Functional Characterization of Monoterpenoid Indole Alkaloid (MIA) Biosynthetic Genes in *Catharanthus roseus*

By

Vonny Salim, B.Sc.

A Thesis

Submitted to the Centre for Biotechnology

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

August, 2013

Brock University

St. Catharines, Ontario

Abstract

The monoterpenoid indole alkaloids (MIAs) of Madagascar periwinkle (*Catharanthus roseus*) are known to be among the most important source of natural drugs used in various cancer chemotherapies. MIAs are derived by combining the iridoid secologanin with tryptamine to form the central precursor strictosidine that is then converted to most known MIAs, such as catharanthine and vindoline that dimerize to form anticancer vinblastine and vincristine. While their assembly is still poorly understood, the complex multistep pathways involved occur in several specialized cell types within leaves that are regulated by developmental and environmental cues.

The organization of MIA pathways is also coupled to secretory mechanisms that allow the accumulation of catharanthine in the waxy leaf surface, separated from vindoline found within leaf cells. While the spatial separation of catharanthine and vindoline provides an explanation for the low levels of dimeric MIAs found in the plants, the secretion of catharanthine to the leaf surface is shown to be part of plant defense mechanisms against fungal infection and insect herbivores.

The transcriptomic databases of *Catharanthus roseus* and various MIA producing plants are facilitating bioinformatic approaches to identify novel MIA biosynthetic genes. Virus-induced gene silencing (VIGS) is being used to screen these candidate genes for their involvement in iridoid biosynthesis pathway, especially in the identification of 7-deoxyloganic acid 7-hydroxylase (CrDL7H) shown by the accumulation of its substrate, 7-deoxyloganic acid and decreased level of secologanin along with catharanthine and vindoline. VIGS can also confirm the biochemical function of genes being identified, such as in the glucosylation of 7-deoxyloganetic acid by CrUGT8 shown by decreased

level of secologanin and MIAs within silenced plants. Silencing of other iridoid biosynthetic genes, loganic acid *O*-methyltransferase (LAMT) and secologanin synthase (SLS) also confirm the metabolic route for iridoid biosynthesis *in planta* through 7-deoxyloganic acid, loganic acid, and loganin intermediates. This route is validated by high substrate specificity of CrUGT8 for 7-deoxyloganetic acid and CrDL7H for 7-deoxyloganic acid. Further localization studies of CrUGT8 and CrDL7H also show that these genes are preferentially expressed within *Catharanthus* leaves rather than in epidermal cells where the last two steps of secologanin biosynthesis occur.

Acknowledgements

First of all, I would like to express my gratitude to Dr. Vincenzo De Luca for the rewarding opportunity to work in his lab for the past five years. His insightful advice, mentorship, and patience are the major contributions in increasing my appreciations of what nature can do.

Secondly, I would like to thank my committee members, Dr. Charles Després and Dr. Art van der Est for their advice throughout my PhD project.

In addition, I am also grateful to the De Luca lab members for my scientific training, especially Dr. Fang Yu, who has taught me molecular biology techniques and Jonathon Roepke for many insightful discussions and teaching me techniques at early stage of my graduate studies.

Lastly, I would like to thank my family and friends in Indonesia, USA, and Canada, who have given me support to have the perseverance in discovering great things in nature and dedications for science.

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Table of Contents

Table of Contents

List of Tables

List of Tables

List of Figures

List of Figures

List of Figures

List of Abbreviations

10-HGO	10-hydroxygeraniol oxidoreductase
16-OMT	16-hydroxytabersonine *O*-methyltransferase
AMV	avian myeloblastosis virus
BEH	ethylene bridged hybrid
BLAST	basic local alignment search tool
C4H	cinnamate 4-hydroxylase
CA	carborundum abrasion
CaMW	cauliflower mosaic virus
cDNA	complementary deoxyribonucleic acid
CMK	4-cytidyl-diphospho-2-C-methyl-D-erythritol kinase
CPR	NADPH: cytochrome P450 reductase
CR	cathenamine reductase
CYP	cytochrome P450
D4H	deacetoxyvindoline 4-hydroxylase
DAD	diode array detection
DAT	deacetylvindoline acetyltransferase
DEPC	diethylpyrocarbonate
DIBOA	2,4-dihydro-1,4-benzoxazin-3-one
DIMBOA	2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one
DL7H	7-deoxyloganic acid 7-hydroxylase
DMAPP	dimethylallyl diphosphate
DXR	1-deoxy-D-xylulose-5-phosphate reductase
DXS	1-deoxy-D-xylulose-5-phosphate synthase
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
EST	expressed sequence tag
EV	empty vector
F3,5H	flavonoid 3',5'-hydroxylase
G10H	geraniol 10-hydroxylase
GES	geraniol synthase
GPP	geraniol diphosphate
GT	glucosyltransferase
HCT	hydrochlorothiazide
HDS	hydroxymethylbutenyl diphosphate synthase
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
HMGR	3-hydroxy-3-methylglutaryl-coenzyme A reductase
HPLC	high-performance liquid chromatography
IPAP	internal phloem-associated parenchyma
IPP	isopentenyl diphosphate
IRS	iridoid synthase
LAMT	loganic acid *O*-methyltransferase
LB	Luria-Bertani
LC-MS	liquid chromatography-mass spectrometry

LS	Linsmaier and Skoog
MAT	minovincinine 19-*O*-acetyltransferase
MCS	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
MCS	multiple cloning sites
MECS	4-cytidyl-diphospho-2-C-methyl-D-erythritol synthase
MeJA	methyl jasmonate
MEP	methylerythritol 4-phosphate
MES	2-(*N*-morpholino) ethanesulfonic acid
MIA	monoterpenoid indole alkaloid
MS	mass spectrometry
MVA	mevalonic acid
MVAPP	mevalonate diphosphate
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NMT	16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase
ODS	octadecylsilyl
ORCA	octadecanoid-responsive *Catharanthus* APETALA2-domain
ORF	open reading frame
PCR	polymerase chain reaction
PDS	phytoene desaturase
POD	peroxidase
PSPG	plant secondary product glycosyltransferase
qRT-PCR	quantitative reverse transcriptase-polymerase chain reaction
RACE-PCR	rapid amplification of cDNA ends-polymerase chain reaction
RNA	ribonucleic acid
RT	retention time
RT-PCR	reverse transcriptase-polymerase chain reaction
SGD	strictosidine β-D-glucosidase
STR	strictosidine synthase
T16H	tabersonine 16-hydroxylase
T19H	tabersonine 19-hydroxylase
TDC	tryptophan decarboxylase
TE	tris- ethylenediaminetetraacetic acid
THAS	tetrahydroalstonine synthase
TLC	thin layer chromatography
TRV	tobacco rattle virus
UDP	uridine diphosphate
UGT	uridine diphosphate-glucosyltransferase
UPLC-MS	ultra-performance liquid chromatography-mass spectrometry
UV/VIS	ultraviolet-visible
VIGS	virus-induced gene silencing

General Introduction

Monoterpenoid indole alkaloids (MIAs) constitute one of the largest and most diverse families of plant secondary metabolites, occurring in Apocynaceae, Loganiaceae, and Rubiaceae. MIA biosynthesis involves the assembly of tryptamine with a secoiridoid, secologanin to yield the central precursor, strictosidine, which can be converted to several thousand biologically active MIAs.

Catharanthus roseus or Madagascar periwinkle is a perennial semishrub of the Apocynaceae family that produces antitumor drugs, vinblastine and vincristine derived from MIAs catharanthine and vindoline. These commercially valuable anticancer dimeric MIAs have been used in chemotherapy treatments against leukemia and other types of cancer. In *C. roseus* plants, these important anticancer agents are produced in very low amounts in the leaves. *C. roseus* also produces hundreds of other MIAs with diverse medicinal properties that are differentially accumulated in roots and aerial organs of the plants. Although *C. roseus* is the most well-studied system for investigating the MIA biosynthesis and some key strategies have been developed for metabolic engineering efforts for increased production of important MIAs by reconstituting the pathway in various biological backgrounds, there are some reactions in the pathway that remain to be elucidated. Functional characterization of genes involved in MIA biosynthesis would provide a better understanding of how the biosynthesis pathway is controlled in plants.

Outline

This thesis contains four chapters. The first chapter is a literature review on recent updates on metabolic pathways of MIA, their regulation and cellular localization, and

technology that can be exploited for functional gene characterization studies with *C. roseus* as a pioneer system for detailed characterization of MIA biosynthesis pathway. The second chapter demonstrates the spatial separation of MIAs, catharanthine and vindoline, in which catharanthine is secreted to the leaf surface while vindole is found within the leaf. The next chapter illustrates the functional characterization of a glucosyltransferase in secologanin biosynthesis with the use of virus-induced gene silencing (VIGS) to confirm the gene function and to clarify the metabolic route of iridoid biosynthesis in *C. roseus*. Lastly, the fourth chapter demonstrates functional characterization of 7-deoxyloganic acid 7-hydroxylase, a cytochrome P450 involved in iridoid biosynthesis of *C. roseus* with the application of VIGS as a tool to find novel genes involved in MIA biosynthesis.

Chapter 1 is a literature review that highlights the complex organization of the MIA pathway in *C. roseus*. This chapter emphasizes the division of the pathways, from the methylerythritol 4-phosphate (MEP), early iridoid, early MIA, to late MIA pathway and describes the use of recent large-scale sequencing databases for functional characterization of genes involved in the MIA biosynthesis and further metabolic engineering efforts. This chapter is a manuscript that has been published in *Advances in Botanical Research* in 2013.

Chapter 2 is a manuscript that has been published in the *Proceedings of the National Academy of Sciences* in 2010. This study addresses the spatial separation of monomers, catharanthine and vindoline which explains the low level production of dimeric MIAs, vinblastine found *in planta*. The metabolite analysis was performed by Vonny Salim [Brock University (BU)]. Enzyme assays were performed by Jonathon

Roepke (BU). The ability of catharanthine to inhibit the growth of fungus on the leaf surface was also tested and its analysis was performed by Vonny Salim (BU). The insect studies were conducted by Vonny Salim (BU) and Kerstin Ploss in Dr. Wilhelm Boland's laboratory at Max Planck Institute for Chemical Ecology in Jena, Germany.

Chapter 3 is a manuscript that has been accepted for publication in *The Plant Cell* in 2013, which describes the functional characterization of a novel glucosyltransferase and the use of VIGS to study secologanin biosynthesis pathway performed by Vonny Salim (BU). The functional characterization of this glucosyltransferase was performed by Keisuke Asada in Dr. Hajime Mizukami's laboratory at Nagoya City University, Japan. The *in-situ* hybridization studies were performed by Elizabeth Edmunds (BU) from the De Luca lab.

Chapter 4 is a manuscript that has been accepted for publication in *The Plant Journal* in 2013 and illustrates the functional characterization of 7-deoxyloganic acid 7-hydroxylase (CrDL7H) involved in secologanin biosynthesis of *C. roseus* guided by bioinformatic analysis of transcriptomic databases generated from large-scale sequencing of various secologanin producing species. The RNA extraction for large-scale sequencing in the De Luca lab was performed by Dr. Sayaka Masada-Atsumi (Apocynaceae species and *Lonicera japonica*), Dylan Levac (BU) (*Rauwolfia serpentina*), and Vonny Salim (BU) (*Cinchona ledgeriana)*. Raw sequences were assembled and annotated by Ye Zhang at University of Calgary under the supervision of Dr. Christoph Sensen. The vector construction for cytochrome P450 expression in yeast was assisted by Dr. Fang Yu, a post-doctoral fellow from the De Luca lab (BU). The cloning, biochemical characterization, and silencing of CrDL7H were performed by Vonny Salim (BU).

Chapter 1 – Literature Review

Towards complete elucidation of monoterpene indole alkaloid biosynthesis pathway: *Catharanthus roseus* as a pioneer system

This manuscript has been published in *Advances in Botanical Research* (2013) 68, 1-37

Towards complete elucidation of monoterpene indole alkaloid biosynthesis pathway: *Catharanthus roseus* as a pioneer system

Vonny Salim and Vincenzo De Luca

Department of Biological Sciences, Brock University, St. Catharines, Ontario, Canada

1.1 Abstract

The development of various plant-based engineering efforts has been facilitated by recent large-scale transcriptomic resources. In consideration of the progress in the study of monoterpene indole alkaloid (MIA) metabolism achieved in the last decade, some strategies have been developed for metabolic engineering efforts. However, unidentified biosynthetic genes in the pathway limit this potential. *Catharanthus roseus* is the most well-studied medicinal plant owing to its production of valuable anticancer dimeric MIAs such as vinblastine. This chapter highlights the cell-, organ-, development-, and environment-specific organization of MIA biosynthesis and describes the intra-and inter-cellular trafficking of MIAs required for their assembly within *C. roseus*. The combined use of cell- and organ-specific transcriptome databases of several MIA-accumulating plants is facilitating bioinformatic approaches to identify MIA candidate genes. Virus-induced gene silencing (VIGS) is being used to screen candidate genes for their involvement in MIA biosynthesis, and the function of selected genes can be identified by the expression and assay of recombinant proteins in bacterial or yeast systems. These new tools show great promise for a more rapid discovery of new genes

involved in whole MIA pathways that enhance the potential of reconstituting them in heterologous microorganisms for the production of any valuable MIA.

1.2 Introduction

Monoterpene indole alkaloids (MIAs) are one of the most diverse groups of plant secondary metabolites of the Apocynaceae, Loganiaceae, and Rubiaceae plant families. MIAs comprise approximately 3000 compounds that exhibit powerful biological activities. While the roles of some MIAs have been described in plant defense against herbivores and pathogens (Luijendijk, Van der Meijden, and Verpoorte, 1996), key derivatives have been exploited for therapeutic purposes. Several drugs have been developed including anticancer agents such as vinblastine and vincristine from *Catharanthus roseus* (Madagascar periwinkle) and camptothecin from *Camptotheca acuminata*, antiarrhythmic agents such as ajmaline, agents for the treatment of neurological disorders such as serpentine from *Rauwolfia serpentina*, and antimalarial agents such as quinine from *Cinchona ledgeriana* (Figure 1.1). Among the best-characterized plants investigated, the MIA biosynthesis pathways from *C. roseus* have been extensively studied at the biochemical and molecular levels (O'Connor and Maresh, 2006; Ziegler and Facchini, 2008). While the Madagascar periwinkle remains the only commercial source for vinblastine and vincristine, this plant also accumulates many different types of MIAs (O'Connor and Maresh, 2006; van der Heijden *et al*., 2004) in addition to these potent inhibitors of microtubule formation that have been developed to treat leukemia, Hodgkin's lymphoma, and other types of cancer.

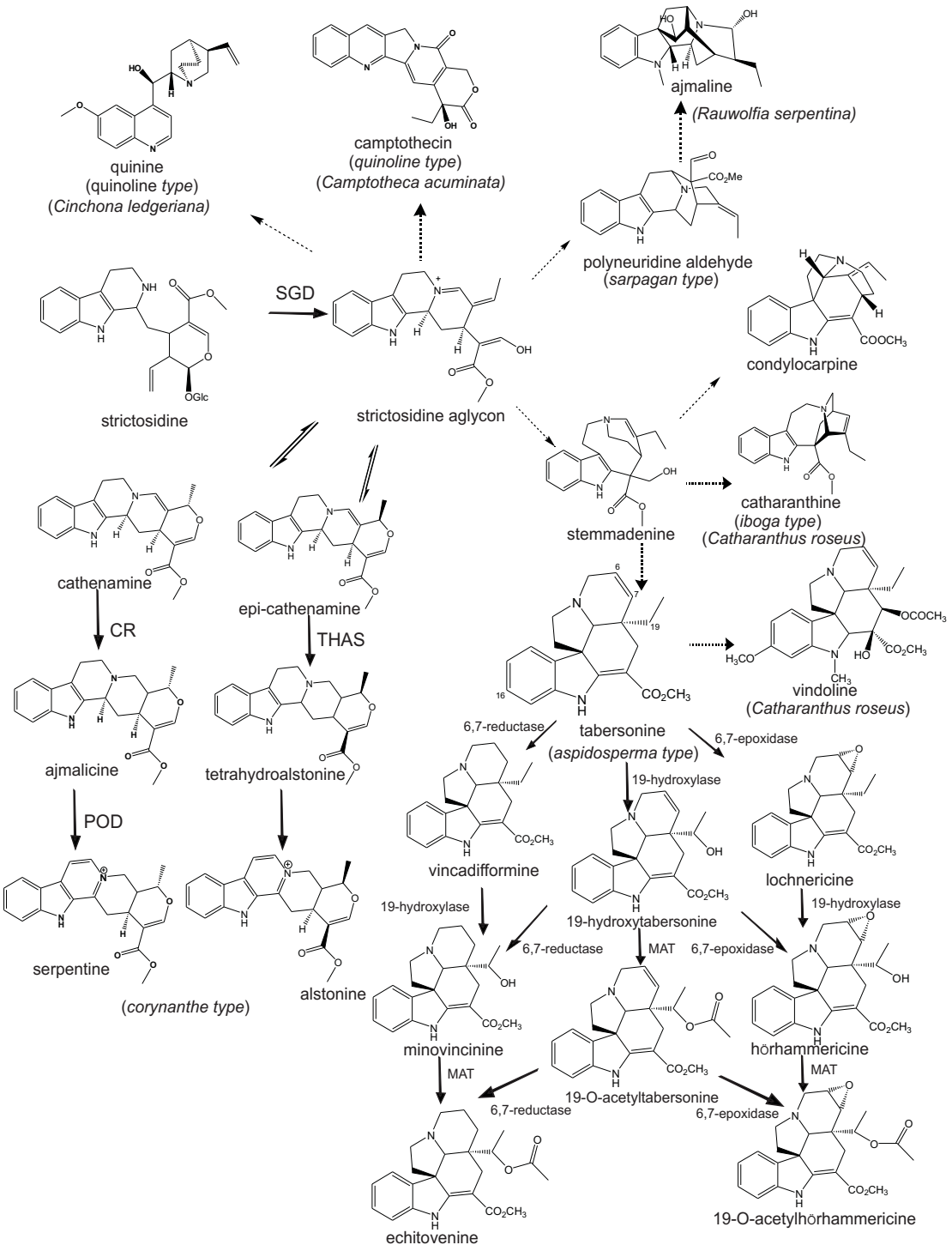


Figure 1.1 The representation of chemical structures of different types of alkaloids derived from strictosidine, namely iboga (catharanthine), aspidosperma (tabersonine that is elaborated into separate pathways to produce vindoline in *Catharanthus* leaves, or MIAs such as lochnericine and hörhammericine in roots), and corynanthe (ajmalicine and serpentine in *Catharanthus*), sarpagan types in *Rauwolfia serpentina*, quinoline types such as camptothecin in *Camptotheca acuminata* and quinine in *Cinchona ledgeriana*. Dashed arrows show multiple steps in biosynthetic pathway. Abbreviations: SGD-strictosidine β-D-glucosidase; CR-cathenamine reductase; THAS-tetrahydroalstonine synthase, POD-peroxidase, MAT- minovincinine 19-*O*-acetyltransferase. *Adapted from El-Sayed and Verpoorte (2007) and Giddings et al., (2011)*.

The early stages of MIA biosynthesis in *C. roseus* involve the formation of the iridoid-secologanin derived from isoprenoid biosynthesis and its condensation with tryptamine to yield the central intermediate, strictosidine, the common precursor for highly divergent MIAs (Figure 1.1) that include catharanthine and vindoline. The two monomers are then coupled to form anticancer dimeric MIA, vinblastine (Costa *et al*., 2008). Despite the efforts by synthetic chemists to produce these valuable secondary metabolites (Ishikawa, Colby, and Boger, 2008; Kuboyama *et al*., 2004), their industrial production still depends on extraction and purification from *C. roseus* leaves with low yields (Gueritte *et al*., 1980). Therefore, an alternative method to improve the production of these valuable molecules would be advantageous. Some efforts will include the transfer of biosynthetic pathways identified in *C. roseus* or other MIA producing plants to microorganisms or to other plant species, which will require enormous technological breakthroughs to be realized. In order to contemplate such large-scale metabolic pathway, engineering the identification and characterization of the remaining biosynthetic genes and other proteins involved in the target pathway is necessary. In addition, MIA biosynthesis in *C. roseus* has been demonstrated to occur in distinct cell types. The translocation of intermediates between cells requires the identification of the transporters involved and shows the additional complexities of MIA biosynthesis that may require understanding and characterization in order to facilitate metabolic engineering of the pathway.

Although a number of genes involved in the biosynthesis of catharanthine and vindoline in *C. roseus* have been identified by traditional forward genetic approaches (enzyme isolation from the plants, protein purification followed by sequencing and

recombinant protein expression in appropriate hosts) many genes remain to be elucidated (De Luca *et al*., 2012b). Homology-based cloning approaches that use sequence similarity to help identify other gene family members, such as acetyltransferase and cytochrome P450 genes, involved in the pathway have been applied. However, this method relies on testing many candidate genes for functional analysis through tedious processes and is often limited by the availability of the substrate for the reaction being identified. Alternatively, the plant can be stimulated with an elicitor to up-regulate the likely MIA genes involved that lead to increases in MIA production (Ziegler and Facchini, 2008). In the last decade, modern genomic and molecular biology methods have begun to accelerate the discovery of MIA biosynthesis steps (De Luca *et al*., 2012a; Facchini *et al*., 2012; O'Connor, 2012; Góngora-Castillo *et al*., 2012; Xiao *et al*., 2013) through the application of high-throughput technologies, including expressed sequence tags (ESTs), DNA microarray analyses, proteomics and metabolomics. Large-scale medicinal plants genome projects such as PhytoMetaSyn (Facchini *et al*., 2012; Xiao *et al*., 2013), and the Medicinal Plant Genomics Consortium [http://www.medicinalplantgenomics.msu.edu/]; (Góngora-Castillo *et al*., 2012) have included studies with *C. roseus* and candidate genes for the remaining steps in MIA biosynthesis in this species are likely to be revealed in the next few years.

This chapter focuses on the biochemistry of MIA biosynthesis, its cell- and organ-specific localization and its regulation by developmental and environmental cues, together with the intra- and inter-cellular trafficking of biosynthetic intermediates required for elaboration of MIAs in the *C. roseus* model system. The chapter also

describes how database mining together with virus-induced gene silencing (VIGS) is being used to speed up the discovery of new MIA genes in *C. roseus*.

1.3 Division of MIA biosynthesis pathway

MIA biosynthetic pathways will be illustrated in four different stages: early steps in monoterpene biosynthesis, iridoid biosynthesis, early MIA biosynthesis, and late MIA biosynthesis emphasizing the vindoline pathway in *C. roseus*.

1.3.1 Early monoterpene biosynthesis

1.3.1.1 Biosynthetic genes involved in the early monoterpene pathway

Terpenes are the largest class of secondary metabolites with over 30,000 compounds derived from C_5 isoprenoid units. These isoprene units condense to form C_{5x} moieties such as monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), triterpenes (C_{30}), tetraterpenes (C_{40}), and polyterpenes (C_{5x}). Terpenes are known to have many biological and physiological functions that affect the normal growth and development of plants as they are the precursors of chlorophyll, and of hormones such as cytokinins, gibberellins, abscisic acid and brassinosteroids (Rodriguez-Conception and Boronat, 2002). The early steps in the isoprenoid pathway consist of the enzymatic steps involved in the synthesis of isopentenyl diphosphate (IPP). Within plants, the biosynthesis of IPP can occur via two metabolic pathways; the first known as the mevalonic acid (MVA) pathway was discovered in the 1950's and occurs in animals, plants, fungi, and some bacteria; the second is the methylerythritol 4-phosphate (MEP) pathway discovered in the mid-1990's and found in most bacteria and plants, but is absent in archaebacteria, fungi, and animals (Rodriguez-Conception and Boronat, 2002; Rohmer, 1999) (Figure 1.2).

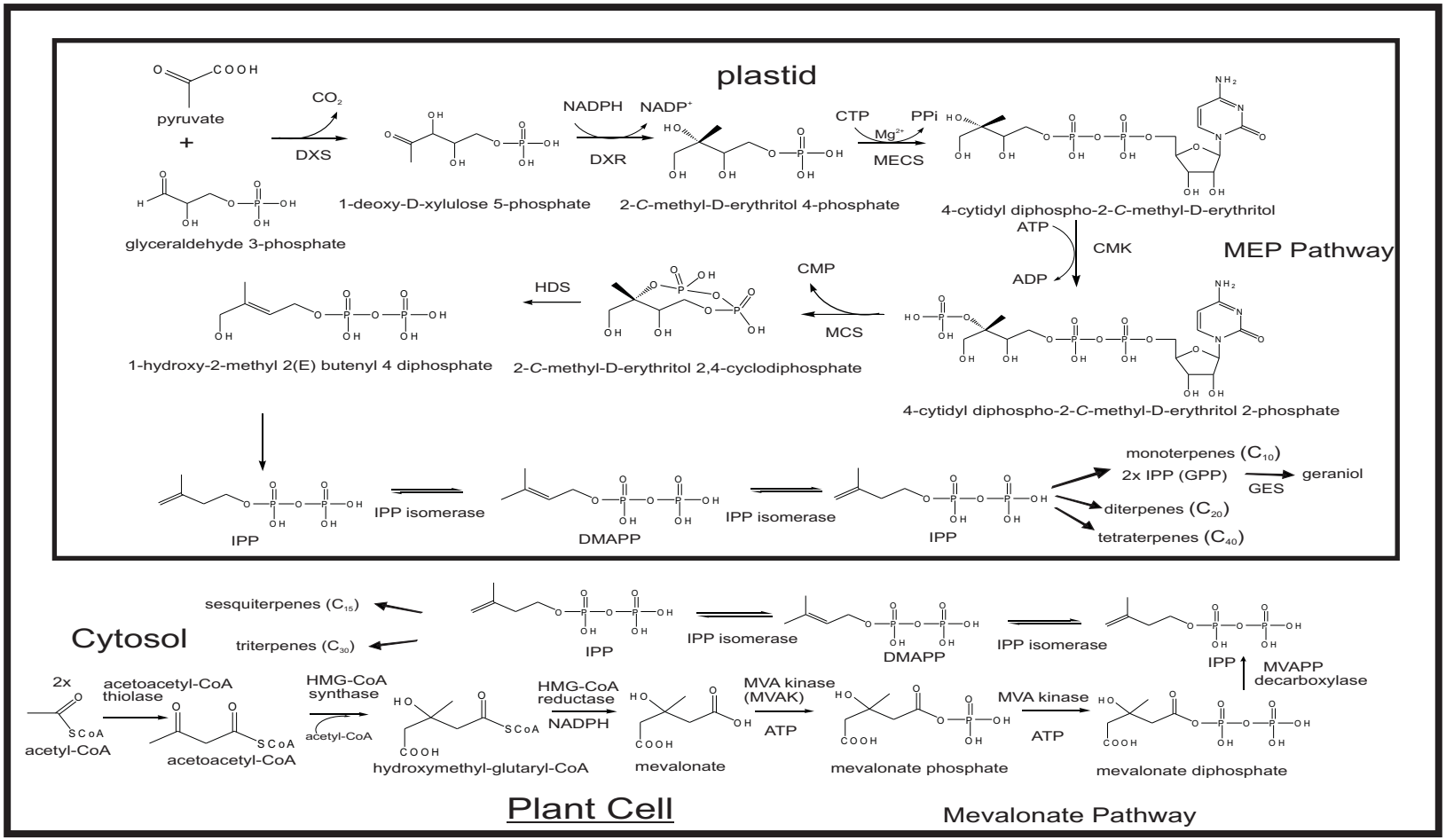


Figure 1.2 The biosynthetic steps of early monoterpene pathway that involves plastidic MEP pathway. Diagram illustrating the subcellular spatial separation and the source of the plant-derived terpenes. Abbreviations: MEP-2-C-methyl-D-erythritol-4-phosphate; DXS-1-deoxy-D-xylulose-5-phosphate-synthase; DXR-1-deoxy-D-xylulose-5-phosphate reductase; MECS-4-cytidyl-diphospho-2-C-methyl-D-erythritol synthase; CMK-4-cytidyl diphospho-2-C-methyl-D-erythritol kinase; MCS-2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS- hydroxymethylbutenyl diphosphate synthase; IPP-Isopentenyl diphosphate; DMAPP-dimethylallyl diphosphate; GES-Geraniol synthase; GPP-Geraniol diphosphate; HMG-CoA-hydroxymethyl-glutaryl-CoA; MVA-mevalonate; MVAPP-mevalonate diphosphate. *Adapted from El-Sayed and Verpoorte (2007)*.

The mevalonate pathway starts with the coupling of two molecules of acetyl-CoA to form acetoacetyl CoA catalyzed by acetoacetyl-CoA thiolase. Condensation of acetoacetyl-CoA with another unit of acetyl-CoA to form 3-hydroxy-3-methyl-glutaryl-CoA (HMG-CoA) is then catalyzed by HMG-CoA synthase. This intermediate is then reduced to form mevalonic acid by HMG-CoA reductase, which is subsequently phosphorylated to produce 5-diphosphomevalonate by mevalonate kinase (MVAK). 5-Diphosphomevalonate is then decarboxylated by 5-diphosphomevalonate decarboxylase to IPP. While MVA pathway, found within the cytosol, provides isoprene units for the assembly of sesquiterpenes and triterpenes (Lange and Croteau, 1999; Newman and Chappell, 1999), the plastid-localized MEP pathway leads to the synthesis of monoterpenes, diterpenes, and tetraterpenes as shown in Figure 1.2. It has been suggested that the biosynthesis of both pathways is highly regulated with cross-talk between the two pathways across the cytosolic and plastid compartments. (Eisenreich, Rohdich, and Bacher, 2001; El-Sayed and Verpoorte, 2007; Oudin *et al*., 2007a).

The initial step of the MEP pathway involves glyceraldehyde 3-phosphate and pyruvate that condense to form 1-deoxy-D-xylulose 5-phosphate (DXP). cDNA encoding 1-deoxy-D-xylulose-5-phosphate synthase (DXS), part of a family of transketolases from *C. roseus*, has been isolated and characterized (Chahed *et al*., 2000). This metabolite is then reduced and isomerized to produce 2-C-methyl-D-erythritol -4-phosphate by DXP reductoisomerase (DXR). In the later step, this intermediate is then condensed with CTP to generate 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol by 4-cytidyl-diphospho-2-C-methyl-D-erythritol synthase (MECS). Both DXR and MECS have been isolated from *C. roseus* and shown to be up-regulated in MIA producing cell cultures (Rohdich *et al*.,

2001; Veau *et al*., 2000) This intermediate is then phosphorylated by ATP to form 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol by a kinase (CMK), and the cytidine nucleotide is removed to form 2-C-methyl-D-erythritol-2,4-cyclodiphosphate by 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS). This intermediate is further dehydrated and reduced to form IPP (El-Sayed and Verpoorte, 2007; Oudin *et al*., 2007b; Rodriquez-Conception and Boronat, 2002; Rohdich *et al*., 2001; Rohmer, 1999). This later step has been functionally characterized in *C. roseus*, known as hydroxymethylbutenyl diphosphate synthase (HDS) (Oudin *et al*., 2007b).

Isomerization of IPP to form dimethylallyl diphosphate (DMAPP) is catalyzed by IPP isomerase as a key step in isoprenoid biosynthesis (Ramos-Valdivia, van der Heijden, and Verpoorte, 1997) (Figure 1.2). DMAPP is condensed with one IPP to form geranyl diphosphate (GPP) which is the precursor for the monoterpenes (Contin *et al.,* 1998). Interestingly, analysis of leaf-epidermis-enriched cDNA libraries has identified 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase (MECS) along with four MVA pathway genes 3-hydroxy-3-methyl-glutaryl CoA reductase (HMGR), 3-ketoacyl CoA thiolase, acetoacetyl-CoA thiolase and HMG-CoA synthase, and three genes common to the MEP/MVA pathways, namely isopentenyl diphosphate isomerase, farnesyl diphosphate synthase (FPPS), and geranyl diphosphate synthase (GPPS) (Murata *et al*., 2008). Recently, geraniol synthase (GES) that catalyzes the conversion of geranyl diphosphate (GPP) to geraniol has been cloned and characterized from *C. roseus* (Simkin *et al*., 2013).

Since the discovery of the MEP pathway in higher plants, the metabolic source of the terpene moiety has been re-established. Contin *et al*. (1998) performed feeding

studies with ^{13}C glucose and showed that the terpenoid moiety of secologanin in cell suspension cultures of *C. roseus* is not MVA derived, but is clearly formed from the MEP pathway. The consistent results of feeding studies with cultures *Ophiorrhiza pumila* also suggest the utilization of the MEP pathway in secologanin biosynthesis (Yamazaki *et al*., 2003). Furthermore, transcript analysis by both Veau *et al*. (2000) and Chahed *et al*. (2000) showed that the expression of the MEP pathway genes is more highly correlated with the accumulation of the MIA ajmalicine, rather than the expression of the MVA pathway genes. Based on these results, MEP pathway derived secologanin is consistent with the known origins of monoterpenes (Eisenreich, Rohdich, and Bacher, 2001).

1.3.1.2 Localization of early monoterpene biosynthesis pathway

The biosynthesis of monoterpenes in plants has been shown to occur in plastids, where the MEP pathway is localized (Lange and Croteau, 1999; Rohmer, 1999). Furthermore, *in situ* hybridization studies have suggested that three genes from the MEP pathway (1-*deoxy-D-xylulose 5-phosphate synthase* (*dxs*), *1-deoxy-D-xylulose 5-phosphate reductoisomerase* (*dxr*), *2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase* (*mecs*) are preferentially expressed in the internal phloem-associated parenchyma (IPAP) cells of the vasculature (Burlat *et al*., 2004). Based on these data, IPAP cells seem to play important roles in providing the precursors for MIA biosynthesis. Immunogold labeling studies of hydroxymethylbutenyl diphosphate synthase (HDS) (Oudin *et al*., 2007b) localized this MEP pathway protein to IPAP, mesophyll, and epidermal cells. However the signal found in IPAP cells was 100- to 250-fold greater than in epidermal cells. Recent studies to characterize geraniol synthase also localized its expression to IPAP cell plastids

(Simkin *et al*., 2013). Together these and other studies strongly suggest that the entire MEP pathway is preferentially expressed in IPAP cells of *Catharanthus* leaves.

1.3.1.3 Gene regulation of early monoterpene biosynthesis

Studies with various tissues such as seedlings and plant organs at various developmental stage, hairy root cultures, and organ cultures suggest that MIA biosynthesis is developmentally regulated (El-Sayed and Verpoorte, 2007; Memelink and Gantet, 2007). For example, a root-specific transcription factor, CrWRKY1 has been studied. Over-expression of CrWRKY1 in roots resulted in the accumulation of serpentine in this specific organ (Suttipanta *et al*., 2011). Furthermore, coordinated expression of genes involved in the pathway is important for optimized metabolite production that is required for plant response to its environment. Methyl jasmonate (MeJA) is one of the signaling molecules that have been studied for triggering defense against pathogens and herbivores. MeJA appears to trigger a signal transduction cascade that activates the octadecanoid-derivative responsive *Catharanthus* APETALA-3 binding protein known as ORCA3. The MeJA mediated transcription factor cascade includes involvement of the basic helix-loop-helix protein, CrMYC2 (Zhang *et al*., 2011) that appears to regulate ORCA3. Cell suspension cultures of *C. roseus* that over-express ORCA3 up-regulated expression of *1-deoxy-D-xylulose 5-phosphate synthase* (*dxs*) gene (El-Sayed and Verpoorte, 2007; van der Fits and Memelink, 2000). *Catharanthus* gene profiling data also showed that additional genes were up-regulated in response to MeJA treatment (*1-deoxy-D-xylulose 5-phosphate synthase* (*dxs*), *2-C-methyl-D-erythritol 2,4,-cyclodiphosphate synthase* (*mecs*), *hydroxymethylbutenyl diphosphate synthase* (*HDS*),

and *geranyl pyrophosphate synthase* (*gpps*)) (Rischer *et al*., 2006). Although this data supports the MeJA mediated up-regulation of the MEP pathway, little information is available on the fate of the IPP produced as a result of this induction for the formation of various terpenes and MIAs.

1.3.2 Iridoid biosynthesis

1.3.2.1 Biosynthetic genes involved in iridoid pathway

The first step to iridoid biosynthesis oxidizes geraniol to generate 10-hydroxygeraniol by a cytochrome P450 monooxygenase (CYP76B6), which was purified cloned and functionally characterized from *C. roseus* (Collu *et al*., 2001; Meehan and Coscia, 1973) (Figure 1.3). 10-Hydroxygeraniol is further oxidized into the dialdehyde 10-oxogeranial by an oxido-reductase (Ikeda *et al*., 1991). 10-Oxogeraniol is converted to iridodial by cyclization (Sanchez-Iturbe, Galaz-Avalos, and Loyola-Vargas, 2005; Uesato *et al*., 1987). Recently, this NADPH-dependent cyclase, iridoid synthase (IRS) has been cloned and characterized from *C. roseus* (Geu-Flores *et al*., 2012) (Figure 1.3). The cyclized product is then further oxidized to form 7-deoxyloganetic acid that involves a cytochrome P450 enzyme followed by glucosylation (GT) to produce 7-deoxyloganic acid and then further hydroxylated (DL7H) to form loganic acid (Madyastha *et al*., 1973) (Figure 1.3). Methylation of loganic acid to form loganin is catalyzed by loganic acid methyltransferase (LAMT) that has been cloned and characterized from *C. roseus* (Murata *et al*., 2008). Finally, the cleavage of the cyclopentane ring of loganin to secologanin by secologanin synthase (SLS) is of particular interest since it is one of the unusual P450s involved in secondary metabolism (Mizutani and Sato, 2010). This

enzyme first detected in a cell suspension culture extracts of *Lonicera japonica* (Yamamoto *et al*., 2000) was suggested to be membrane-associated and to be a P450 with requirements for NADPH and O_2. Molecular cloning of SLS (CYP72A1) from *C. roseus* was followed by its functional expression and biochemical characterization in *Escherichia coli* confirmed its ability to convert loganin into secologanin (Irmler *et al*., 2000) (Figure 1.3).

1.3.2.2 Localization of the iridoid pathway

The enzymes of iridoid pathway are localized to separate subcellular compartments and involve more than one cell-type. As geraniol synthase has been localized to the plastid, geraniol may then be exported to the cytosol where it would be converted to 10-OH geraniol by G10H associated with vacuolar membranes, and later with the endoplasmic reticulum (Madyastha *et al*., 1977; Guirimand, *et al*., 2009) and is associated with membrane-bound flavin containing NADPH: cytochrome P450 reductase (CPR) normally required for cytochrome P450 reactions. In addition to localization of three identified MEP pathway genes by *in situ* hybridization, Burlat *et al*. (2004) also showed that G10H is preferentially expressed in the IPAP cells. Very recent studies with iridoid synthase (IRS) also localized expression of this gene to IPAP cells (Geu-Flores *et al*., 2012), while the 2nd to last (Roepke *et al*., 2010) and the last steps (Irmler *et al*., 2000) in secologanin biosynthesis appear to be localized to epidermal cells. Together these data suggest that IPAP cells are specialized to supply the monoterpene, geraniol for the iridoid pathway, at least up to the cyclization step. While the cellular localization of genes involved the formation of the carboxyl group, glucosylation and hydroxylation to

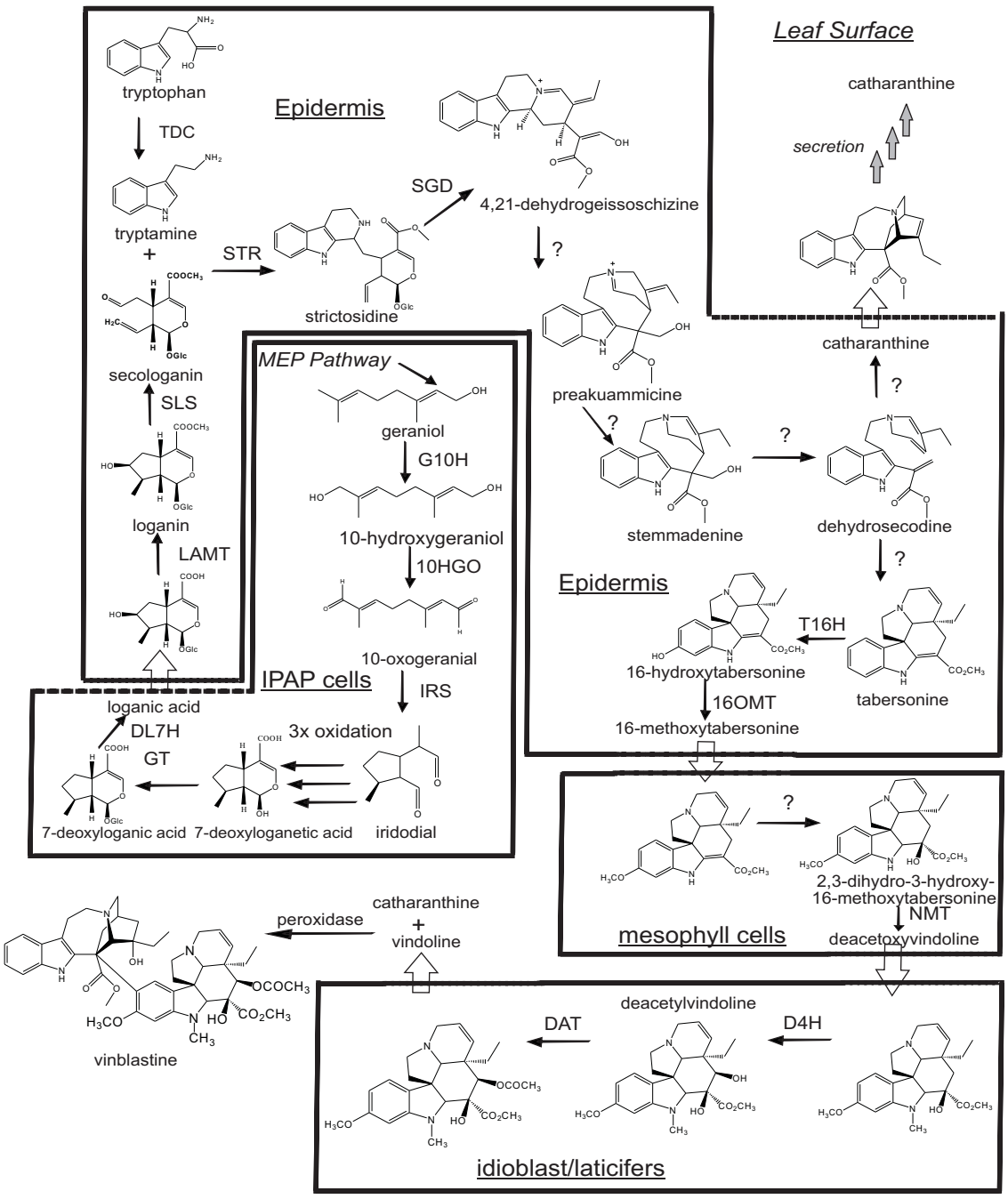


Figure 1.3 The MIA biosynthetic pathway is compartmentalized in different *Catharanthus* leaf cells. The early iridoid pathway is localized to the internal phloem associated parenchyma (IPAP) cells while the late iridoid pathway and most of the MIA biosynthesis are associated with leaf epidermis, and the late vindoline pathway is localized to special idioblast/laticifer cells. Open arrows show transport of metabolites. Abbreviations: TDC-tryptophan decarboxylase; STR-strictosidine synthase; SGD-strictosidine-β-D-glucosidase; G10H-geraniol 10-hydroxylase; 10HGO-10-hydroxygeraniol oxidoreductase; IRS-iridoid synthase; GT-glucosyltransferase; DL7H-7-deoxyloganic acid 7-hydroxylase; LAMT-loganic acid –*O*-methyltransferase; SLS-secologanin synthase; T16H-tabersonine 16-hydroxylase; 16OMT-16-hydroxytabersonine-*O*-methyltransferase; NMT-*N*-methyltransferase; D4H-deacetoxyvindoline 4-hydroxylase; DAT-deacetylvindoline acetyltransferase. '?' indicate genes that have not been cloned and biochemically characterized. *Adapted from Roepke et al., (2010)*.

produce loganic acid have yet to be determined; it is likely that they may also occur in IPAP cells (Figure 1.3). If this is correct, an undetermined loganic acid transporter may then be involved in its export from IPAP cells to the epidermis for final elaboration into secologanin.

1.3.2.3 Gene regulation of iridoid biosynthesis

While few data are available, there seem to be similarities in the regulation of the early steps of iridoid biosynthesis by a MeJA signaling pathway. Gene profiling data showed that expression of G10H and 10HGO was significantly higher in cells treated with MeJA, while the terminal leaf epidermis localized SLS expression was not affected (Rischer *et al*., 2006). While these studies may indicate that the early steps of iridoid biosynthesis are regulated differently than the later steps, it is not clear if the spatial separation of this pathway may also play a role in this process.

1.3.3 Early MIA biosynthesis

1.3.3.1 Biosynthetic genes involved in the early MIA pathway

The central precursor strictosidine is formed by a stereoselective Pictet-Spengler condensation mediated by strictosidine synthase (STR) of tryptamine derived from tryptophan by a pyridoxal-phosphate dependent tryptophan decarboxylase (TDC) (De Luca, Marineau, and Brisson, 1989) and secologanin (Figure 1.3) (de Waal, Meijer and Verpoorte, 1995; Maresh *et al*., 2008; Stöckigt and Zenk, 1977; Treimer and Zenk, 1979). TDC has been cloned and functionally characterized from different MIA-producing plants (De Luca, Marineau, and Brisson, 1989; Lopez-Meyer and Nessler, 1997;

Yamazaki *et al*., 2003). Strictosidine synthase (STR) was first cloned from *R. serpentina* and functionally expressed in *E. coli* (Kutchan *et al*., 1988). Later, STR orthologs were isolated from *C. roseus* and *O. pumila* (McKnight *et al*., 1990; Pasquali *et al*., 1992; Yamazaki *et al*., 2003). The glucose moiety of strictosidine is subsequently removed by strictosidine β-D-glucosidase (SGD) (Figure 1.3). Further purification of SGD from *C. roseus* cell cultures revealed that SGD has a strong affinity for strictosidine as substrate and a high molecular mass that exists as an aggregate of multiple 63-kDa subunits (Luijendijk, Stevens, and Verpoorte, 1998). While trypsin digestion was performed to solubilize the enzyme without the loss of activity, this stability to proteases has been used to suggest a putative but uncharacterized role for this enzyme in plant defense. SGD was later functionally characterized from *C. roseus* and *R. serpentina* (Geerlings *et al*., 2000; Gerasimenko *et al*., 2002). Encoded by a single gene in *C. roseus*, SGD shares about 60% homology at the amino acid level with other plant glucosidases (Geerlings *et al*., 2000). The three dimensional structure of the SGD enzyme was also studied to reveal its catalytic mechanism (Barleben *et al*., 2007). Interestingly, the production of a versatile strictosidine aglycone is the driving force combined with a number of different uncharacterized enzyme-mediated reactions for the remarkable structural diversity of MIAs found in nature.

The ring arrangements after the formation of this reactive hemiacetal intermediate seems to be species dependent (Szabo, 2008; Zhu *et al*., 1990). These diverse metabolites are produced only in certain plant families (Apocynaceae, Loganiaceae, and Rubiaceae) and each member produces a subset of compounds that contributes to its varying biological function (Szabo, 2008). Within *C. roseus*, different arrangements of

strictosidine aglycone yield three major (corynanthe, iboga, and aspidosperma) classes of MIAs (Qureshi and Scott, 1968) (Figure 1.1).

Downstream steps of strictosidine aglycone formation leading to ajmalicine have been partially characterized. Following the conversion of strictosidine by *C. roseus* SGD, carbinolamine serves as an intermediate to produce cathenamine that is then reduced to form ajmalicine by cathenamine reductase (CR). Two different CRs with requirements for NADPH as cofactors have been identified and are detected at low levels in *C. roseus* cell cultures (El-Sayed and Verpoorte, 2007); one reduces cathenamine into ajmalicine and 19-epi-ajmalicine while the other converts the iminium form of cathenamine into tetrahydroalstonine by tetrahydroalstonine synthase (THAS) (Hemscheidt and Zenk, 1985). Later, ajmalicine can be converted into serpentine by a vacuolar peroxidase (POD) from *C. roseus* (Blom *et al*., 1991) (Figure 1.1).

The preakummicine, strychnos-type intermediate may be the common precursor for the aspidosperma, strychnos, and iboga alkaloids (Figure 1.1). Although mechanisms of preakuammicine formation from strictosidine aglycon have been proposed, preakummicine has not been detected in plant extracts, mainly due to its lability. Moreover, preakuammicine can be reduced to stemmadenine, a more stable intermediate that can be rapidly consumed in the cell culture to yield catharanthine and tabersonine skeletons, as well as condylocarpine, but not as an intermediate (El-Sayed *et al*., 2004) (Figure 1.1). Stemmadenine may be rearranged to generate dehydrosecodine that can lead to catharanthine type alkaloids by a Diels-Alder reaction (Qureshi and Scott, 1968; O'Connor and Maresh, 2006).

Although genes involved in the pathways leading to catharanthine and tabersonine have not been fully isolated, some branch pathways through tabersonine in *C. roseus* and polyneuridine aldehyde have been studied most thoroughly. The six steps that catalyze the conversion of tabersonine to vindoline have been described in detail (Liscombe, Usera, and O'Connor, 2010; Ziegler and Facchini, 2008).

1.3.3.2 Localization of early MIA biosynthesis

While at least three cell types appear to be required to elaborate MIAs (Figure 1.3), this biosynthetic pathway also requires the coordination of IPAP cell-based loganic acid biosynthesis with its transport to the leaf epidermis for conversion to secologanin and with the early MIA pathway. *In situ* hybridization and immunological studies have localized tryptophan decarboxylase (*tdc*) transcript and antigen as well as strictosidine synthase (*str*) transcript to the epidermis of stems, leaves, flower buds, and most protoderm and cortical cells of the apical meristems in root tips (St-Pierre, Vázquez-Flota, and De Luca, 1999). Moreover, RT-PCR of laser-capture microdissected cells showed that *tdc*, *str,* and strictosidine β-D-glucosidase (*sgd*) were preferentially expressed in the epidermis of *C. roseus* (Murata and De Luca, 2005). These data provide strong evidence that leaf epidermal cells are biosynthetically active sites for early MIA biosynthesis. Within epidermal cells, it seems that the first committed step of MIA biosynthesis catalyzed by strictosidine synthase (STR) is localized to vacuoles (McKnight *et al*., 1991; Stevens, Blom, and Verpoorte, 1993), suggesting that tryptamine and secologanin are imported through the vacuole membrane from the cytosol where these metabolites are biosynthesized (De Luca and Cutler, 1987). Strictosidine must then be exported out of the

vacuole into the cytosol for deglucosylation by SGD suggested to be associated with the endoplasmic reticulum as the SGD protein also has a putative targeting signal peptide (Geerlings *et al*., 2000). However, more recent studies have provided intriguing results reporting that SGD may mostly be associated with the nucleus (Guirimand *et al*., 2010). While this result is difficult to reconcile with the involvement of SGD with MIA biosynthesis, it suggests that elaboration of downstream MIAs from strictosidine is likely to be significantly more complex than originally anticipated. While early studies (Hemscheidt and Zenk, 1985) have suggested that separate tetrahydroalstonine, ajmalicine and epi-ajmalicine synthase enzymes are involved in the biosynthesis of these three MIAs, the correct recombinant reductases have yet to be isolated and functionally characterized. Studies with plant vacuoles have further suggested the existence of highly efficient transporters that mobilize newly synthesized ajmalicine for oxidation into serpentine that is then trapped in the vacuole by an ion-trap mechanism (Blom *et al*., 1991; El-Sayed and Verpoorte, 2007). The reactivity of the strictosidine aglycone intermediates required for developing enzyme assays together with other unknown constraints appear to make it very difficult to characterize these sections of the MIA biosynthesis pathway.

1.3.3.3 Gene regulation of early MIA biosynthesis

In addition to MeJA, a fungal elicitor such as *Phythium aphanidermatum* is also known to induce the expression of strictosidine synthase (*str*) and tryptophan decarboxylase (*tdc*) (Roewer *et al*., 1992). Over-expression of ORCA3 transcription factor in cell cultures activated expression of *STR*, *TDC,* and *SGD* (van der Fits and

Memelink, 2000). This result was supported by gene profiling data (Rischer *et al*., 2006) except for *STR* whose expression levels did not change. Another transcription factor associated with MIA biosynthesis is *ORCA2* that was identified after a yeast one-hybrid screen using MeJA or elicitor-responsive regulatory elements in the *STR* promoter (Menke *et al*., 1999). Further promoter analysis showed that it contains a G-Box element that is involved in binding the MeJA-responsive transcription factor CrMYC1 (Chatel *et al*., 2003). Other Zinc Cys_2/His_2-type, Zinc-finger transcription factors such as *ZCT1*, *ZCT2*, *ZCT3* have been shown to bind to the promoter of *TDC* and *STR* functioning as repressors of the MeJA response (Pauw *et al*., 2004). These data suggest that early MIA pathway genes (*TDC*, *STR,* and *SGD*) respond to MeJA treatment in a well coordinated manner. It is interesting that the accumulation of catharanthine, tabersonine, and vindoline is also increased in response of MeJA treatment (Aerts *et al*., 1996), but the molecular basis of this observation in not known.

1.3.4 The late MIA biosynthesis pathway

1.3.4.1 Biosynthetic genes involved in late MIA / vindoline pathway in C. roseus

The late steps of MIA biosynthesis in *C. roseus* include the production of the well-studied tabersonine derivatives specific to different plant organs. In root tissues tabersonine is oxidized into hörhammericine and lochnericine while it is converted into vindoline in above ground organs (Figure 1.1) (Laflamme, St-Pierre, and De Luca, 2001). Roots oxidize tabersonine to its epoxide via a microsomal cytochrome P450 mono-oxygenase requiring NADPH and molecular oxygen (tabersonine 6,7 epoxidase) in *C. roseus* hairy root cultures (Rodriguez *et al*., 2003). In addition, the root-tip-specific

minovincinine 19-*O*-acetyltransferase (MAT) gene involved in the formation of 6,7-dehydroechitovenine and/or 19-*O*-acetylhörhammericine has been functionally characterized (Laflamme, St-Pierre, and De Luca, 2001) (Figure 1.1). Another cytochrome P450 monooxygenase involved in this root-specific pathway is recently cloned and characterized CYP71BJ1 that hydroxylates at the C-19 of tabersonine and lochnericine (Giddings *et al*., 2011).

Vindoline, one of the monomers of the bisindole alkaloid vinblastine is derived from tabersonine through six enzymatic steps. The first step involves hydroxylation of tabersonine by tabersonine 16-hydroxylase (T16H), another cytochrome P450 monooxygenase (Schröder *et al*., 1999; St. Pierre and De Luca, 1995). Expression of *T16H* is strongly induced by light with low activity found in etiolated seedlings. The hydroxyl group of 16-hydroxytabersonine is then *O*-methylated by an S-adenosyl-methionine (AdoMet) 16-hydroxytabersonine-16-*O*-methyltransferase (16OMT) that has been cloned and characterized to yield 16-methoxytabersonine (Fahn *et al*., 1985; Levac *et al*., 2008). The next uncharacterized step converts 16-methoxytabersonine to 16-methoxy-2,3-dihydro-3-hydroxytabersonine, which is further *N*-methylated to produce deacetoxyvindoline. This *N*-methyltransferase has been cloned and functionally characterized (Liscombe, Usera, and O'Connor, 2010) while its activity has been localized previously within the thylakoid membrane of chloroplasts (De Luca and Cutler, 1987) and was detected in differentiated plants but not in plant cell cultures (Dethier and De Luca, 1993). Deacetoxyvindoline is hydroxylated by the oxoglutarate-dependent dioxygenase, deacetoxyvindoline 4-hydroxylase (D4H) (De Carolis *et al*., 1990; De Carolis and De Luca, 1993; Vazquez-Flota *et al*., 1997). The last step of vindoline

biosynthesis is catalyzed by deacetylvindoline *O*-acetyltransferase (DAT), part of the BAHD (benzylalcohol-*O*-acetyl-, anthocyanin-*O*-hydroxycinnamoyl-, anthranilate-*N*-hydroxycinnamoyl/benzoyl-, and deacetylvindoline 4-*O*-acetyltransferase) family that is responsible for acetylation of deacetoxyvindoline to yield vindoline (De Luca and Cutler, 1987; St-Pierre *et al*., 1998). Interestingly, DAT shares 78% homology at the amino acid level with minovincinine 19-*O*-acetyltransferase (MAT) (Laflamme, St-Pierre, and De Luca, 2001). While DAT accepts only its natural substrate, MAT shows slight activity toward deacetylvindoline, the DAT substrate. The last two steps in vindoline biosynthesis are light-regulated and only found in differentiated plant materials.

1.3.4.2 Localization of vindoline biosynthesis

It is clear that MIA biosynthesis in *C. roseus* occurs in several different cell types (Guirimand *et al*., 2011b; Murata *et al*, 2008) and is developmentally regulated (De Luca *et al*., 1986; Facchini and De Luca, 2008). Although some transcripts, and activities specific to vindoline biosynthesis have been detected in hairy root and cell suspension cultures, these tissues do not produce vindoline. The conversion of tabersonine to vindoline occurs in various compartments, the first enzyme that acts on tabersonine is associated with the endoplasmic reticulum membrane while the second enzyme, 16-OMT is believed to be cytosolic (St. Pierre and De Luca, 1995). Furthermore, *N*-methyltransferase is believed to be associated with the thylakoid within the chloroplast suggesting that this enzyme is found within a chloroplast-rich part of leaf tissue, the mesophyll layer (De Luca and Cutler, 1987; Dethier and De Luca, 1993; Murata and De Luca, 2005). It has been proposed that 16-methoxytabersonine may be transported from

the leaf epidermis where 16-OMT is found (Levac *et al*., 2008), to other cell types for *N*-methylation, passing through cell walls or via the plasmadesmata. At the subcellular level, the methylated intermediate is then transported to the cytosol for further hydroxylation and acetylation (De Carolis *et al*., 1990; De Luca and Cutler, 1987). RNA blot hybridization studies indicated that the enzyme activity followed the levels of *d4h* transcripts occurring primarily in young leaves that decline with age and lower levels occurred in the stem and fruits. St. Pierre *et al*. (1999) reported that expression of D4H and DAT was localized to laticifer and idioblasts cells of leaves, stems and flower buds. These results explain the failure of vindoline production by cell culture technology as the late steps of vindoline biosynthesis take place in different cells and only in above-ground tissues.

1.3.4.3 Vindoline biosynthesis is modulated by MeJA and light

The vindoline pathway of *C. roseus* can be modulated by application of MeJA or by light to developing seedlings. MeJA treatment increased the expression level of deacetoxyvindoline 4-hydroxylase (D4H) (Vazquez-Flota and De Luca, 1998a; 1998b) during seedling development. Furthermore, over-expression of the ORCA3 transcription factor (van der Fits and Memelink, 2000), but not MeJA treatment (Rischer *et al*., 2006), up-regulated *D4H* expression in cell cultures.

Exposure of *C. roseus* cell cultures to light increased the expression of T16H within 22-28 h of treatment (Schroder *et al*., 1999). Etiolated seedlings exposed to full or red light induced the expression of *D4H* together with enzyme activity and this could be reversed by far-red light treatment (Vazquez-Flota *et al*., 1998a). Further studies

suggested the D4H activity could be post-translationally modified by an uncharacterized phytochrome-assisted mechanism. Similarly, DAT involved in the last step in vindoline biosynthesis appears to be activated by light through the involvement of phytochrome that mediates the reversible activation of *DAT* (Aerts and De Luca, 1992). Studies with purified DAT have shown it to be more strongly inhibited by tabersonine and coenzyme A than by tryptamine, secologanin, and vindoline (Power, Kurz, and De Luca, 1990) and raise questions whether such pathway precursors or co-substrates could modulate the *in vivo* activity of DAT.

1.4 Organization and spatial separation of MIA biosynthesis

While enzymes involved in MIA biosynthesis are associated with various subcellular compartments such as the cytosol, vacuole, endoplasmic reticulum, and chloroplast, the division of MIA biosynthesis between multiple cells deserves particular attention. While little is known about the process, subcellular trafficking of intermediates plays a significant role in controlling MIA biosynthesis to effectively channel iridoids/MIA intermediates between cells and remove toxic MIA end products from the cytosol (Figure 1.3). In this context, the low-level accumulation of dimeric anticancer MIA, vinblastine could be due to the spatial separation catharanthine that is secreted into the leaf exudates and vindoline that likely accumulates within idioblast and laticifer leaf cells (Roepke *et al*., 2010). This suggests that the MIA part of biosynthesis takes place in leaf epidermal cells in order to allow the secretion of catharanthine and vindoline pathway intermediates to their respective locations to prevent the accumulation of dimeric MIAs that are known to be toxic to *C. roseus* plants. In order to understand this

complex compartmentalization, the transporters involved need to be identified in relation with the supply of intermediates from one cell type to another and MIA product accumulation in a separate cell type or location.

This multi-cell-type coordination of MIA biosynthesis in leaves contrasts strongly with the occurrence of the MEP pathway, the secologanin pathway, tryptamine production and the complement of enzymes leading to root-based MIAs within protoderm and cortical cells around the root apical meristem. It is interesting that these root cells preferentially expressed *TDC*, *STR*, and the terminal *MAT* involved in the assembly of 6,7-dehydroechitovenine and/or 19-*O*-acetylhörhammericine (Laflamme, St-Pierre, and De Luca, 2001; St. Pierre, Vázquez-Flota, and De Luca, 1999), while D4H and DAT transcripts were not detected in the underground tissue. Remarkably, few studies have been performed to analyze the reasons for the same cell MIA biosynthesis occurring in roots in relation to the multi-cell MIA biosynthesis occurring in leaves of *Catharanthus*.

1.4.1 Epidermis as important biosynthetic sites of MIAs and their precursors

The involvement of at least three cell types, including IPAP, epidermis and idioblast/laticifers indicates the translocation of at least one pathway intermediate in the MIA biosynthesis. Early studies have shown that TDC and STR are most abundant in *C. roseus* roots even though they are also detected in aerial organs (Pasquali *et al*., 1992). In *situ* hybridization studies imply that the transcripts of early MIA pathway, TDC, SLS, and STR are associated with the epidermis of leaves, stems, and flower buds, while the later steps that involve D4H and DAT transcripts are localized to the specialized laticifer

and idioblast cells of the same organs (De Luca and St-Pierre, 2000; Irmler *et al*., 2000; Vazquez-Flota and De Luca, 1998a; 1998b). Furthermore, gene transcripts encoding enzymes in the MEP pathway and geraniol 10-hydroxylase (G10H) are associated with the IPAP cells of young *C. roseus* aerial organs (Burlat *et al*., 2004). These data suggest the potential translocation of vindoline biosynthetic intermediates from the IPAP cells to leaf epidermis and from the epidermis to the latificifers and idioblast (Figure 1.3). As the epidermis seems to be the centralized site for translocation for the pathway intermediates, efforts towards RNA isolation from this particular cell type to identify more genes involved in the entire pathway can be beneficial. This approach has been applied in a number of cases such as the sequencing of specialized glandular trichome cells of *Mentha piperita* to dissect the menthol and polymethylated flavone biosynthesis pathways (Lange *et al*., 2000; Schilmiller, Last, and Pichersky, 2008). The carborundum abrasion (CA) technique was developed to extract RNA from the leaf epidermis and cDNA libraries were generated to produce ESTs that can be exploited for gene discovery (Murata and De Luca, 2005; Murata *et al*., 2006; Murata *et al*., 2008). In addition, the CA technique can be used to isolate large-scale active epidermal protein from the leaf tissues such as in the purification of 16OMT or for enzyme assays of LAMT in which LAMT activity was enriched in protein extract from leaf epidermis compared with whole-leaves (Levac *et al*., 2008; Murata and De Luca, 2005; Murata *et al*., 2008).

1.4.2 The use of epidermis-enriched transcriptomic resources for gene discovery

While random sequencing of leaf epidermis-enriched cDNA libraries (Murata *et al*., 2006) has assisted in the identification of genes involved in MIA biosynthesis in *C.*

roseus, such as characterization of LAMT by homology-based cloning (Murata *et al*., 2008), the availability of this type of database can be used in comparative bioinformatic studies with the large transcriptome *Catharanthus* database (http://www.phytometasyn.ca/) to identify other novel MIA genes. For example, an interesting candidate gene that is firstly pulled out from the whole leaf EST library can be verified for its preferential expression in the leaf epidermal cells (De Luca *et al*., 2012b). As biosynthetic genes in the catharanthine pathway have not been identified, the leaf epidermis-enriched cDNA library represents an excellent auxiliary source for the cloning and characterization of members of this pathway.

1.5 Large-scale genomic approaches in functional characterization of genes involved in MIA biosynthesis

Large-scale transcriptomic projects have recently opened a new approach for revealing genes associated with MIA biosynthesis. Early EST-based approaches yielded more than 25,000 annotated ESTs based on partial sequencing of cDNA libraries from *C. roseus* organs and tissues (Murata *et al*., 2006; Rischer *et al*., 2006; Shukla *et al*., 2006). The correlation networks that combine transcriptomic and metabolomic data from *C. roseus* treated with MeJA by using cDNA/amplified fragment length polymorphism approach was performed to show a correlation between the biosynthetic gene expression and the subsequent metabolite product in MIA biosynthesis pathway (Rischer *et al*., 2006). In general, the results suggest a coordinated increase in the expression of genes associated with MIA metabolism. Furthermore, the number of sequences related to alkaloid-accumulating plants, such as *C. roseus* continues to increase as sequencing of

various tissues such as stems, roots, flowers, elicited cell cultures, and hairy roots including a TDC silenced line has been announced (Góngora-Castillo *et al*., 2012; Runguphan, Maresh, and O'Connor, 2009). These large-scale transcriptome sequences were generated by the Medicinal Plant Genomics Consortium (http://medicinalplantgenomics.msu.edu) that has been initiated with an effort to utilize next-generation sequencing for transcriptomic analysis of 14 medicinal plants.

1.5.1 The shared pathways among Apocynaceae family

The recent large-scale sequencing project, PhytoMetaSyn have produced transcriptomic profiles for over 70 medicinal plant species (http://www.phytometasyn.ca/) that produce secondary metabolites with high values (Facchini *et al*., 2012; Xiao *et al*., 2013). In this project, ESTs from eight alkaloid producing plants belonging to the order Gentianales with seven Apocynaceae family members that include three species of *Catharanthus* (*roseus*, *longifolius*, and *ovalis*), *Vinca minor*, *Rauwolfia serpentina*, *Tabernaemontana elegans*, *Amsonia hubrichtii* and one member of the Rubiaceae (*Cinchona ledgeriana*) have been obtained. In addition, one species, *Lonicera japonica* (Caprifoliaceae) that produces only secologanin (Kawai, Kuroyanagi, and Ueno, 1988), has been sequenced. While over 56,000 ESTs have been reported from leaf, leaf epidermis, and roots of *C. roseus* as part of the NapGen (Natural Products Genomic Research) project (Murata *et al*., 2006), the PhytoMetaSyn project expands the availability of EST that would enable the transcriptome analysis within the same family. As Apocynaceae members share the same pathways for the biosynthesis of various MIAs, the orthologous genes that have been characterized from *C. roseus* to be

involved in the MEP, iridoid, and early MIA pathway, can be found in other seven species (Figure 1.4).

Similarly, the putative iridoid biosynthetic genes from *Lonicera japonica* can be represented in the database. The approach of sequencing several species from the same family provides the advantage in targeting the genes associated with a certain part of the pathway. As a result, this method will facilitate the gene discovery process.

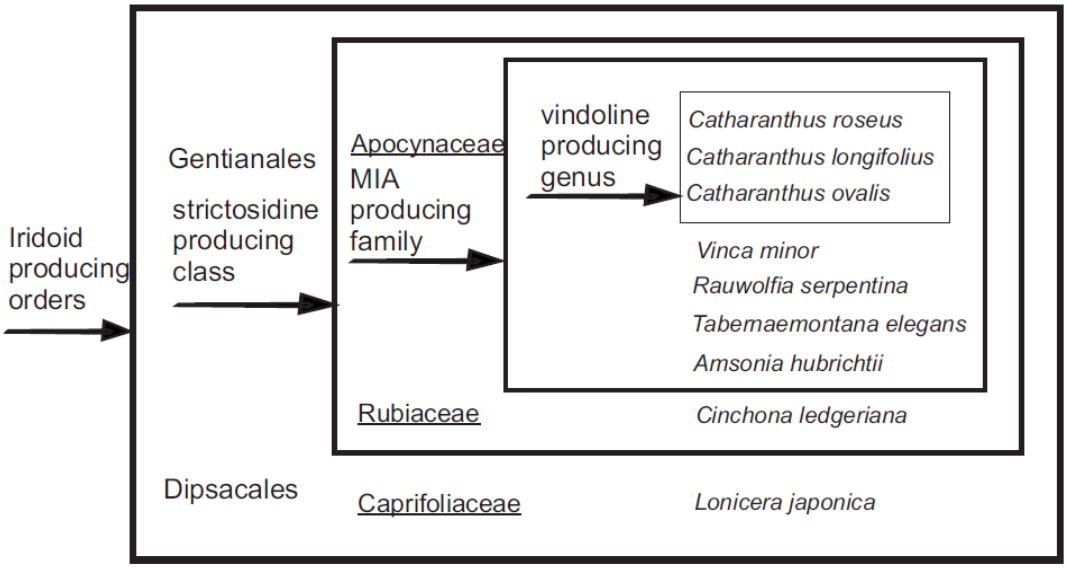


Figure 1.4 Diagram describing an approach for gene discovery by orthology-based comparison among MIA and iridoid-producing plants. The target genes in MEP and iridoid pathways should be found in the databases in commonly iridoid-producing orders, such as Gentianales and one iridoid-only-producing species, *Lonicera japonica* from Caprifoliaceae family, part of Dipsacales order. The target genes for the early MIA pathway can be pulled out from all Gentianales (*Apocynaceae* and *Rubiaceae*) plant databases and those for the late MIA pathway such as vindoline should be uniquely found in the *Catharanthus* databases.

1.5.2 Tools for screening the candidate genes

With the increase in the number of publicly available sequences from MIA producing species, the need for an effective tool to screen a large number of candidate MIA genes has increased. RNA interference (RNAi) is one of the reverse genetics approach that has been used to silence genes in cell suspension cultures (Courdavault *et*

al., 2005; Papon *et al*., 2004) and in hairy root cultures (Runguphan, Maresh, and O'Connor, 2009). RNAi involves the introduction of short double-stranded RNA during plant transformation to interfere with expression of the target gene. Although this technique has been successful in suppressing the production of metabolites in the MIA pathway, experiments that involve RNAi take months or even years before the results become apparent. In addition, this approach is not ideal for characterization of genes in the pathway that shows restricted metabolism in cell culture systems, such as vindoline biosynthesis in *C. roseus*. Furthermore, efficient plant transformation strategies are not always available for most medicinal plant, including *C. roseus*.

Virus-induced gene silencing (VIGS) has become an effective tool for gene functional analysis. This method exploits the plant defense mechanism against virus infection that leads to degradation of the corresponding mRNA transcripts of a host gene that has been targeted for silencing (Burch-Smith *et al*., 2004). The binary tobacco rattle virus (pTRV) vector system developed and tested in *Nicotiana benthamiana* (Dinesh-Kumar *et al*., 2003; Liu *et al*., 2002; Ratcliff, Martin-Hernandez, and Baulcombe, 2001) has also been used with success in other plant systems. The first advantage of VIGS is the rapidity and ease of use of this gene silencing process. The experiments can be performed in *C. roseus* with as few as eight weeks starting from seed germination to screening of individual silenced plants for reduced gene expression and changing metabolite profiles (De Luca *et al*., 2012b). Other convenient advantages include the non-requirement for stable plant transformation and the requirement of only partial sequence information to achieve adequate gene silencing (Senthil-Kumar and Mysore, 2011). Its disadvantages may include non-targeted gene silencing and possible toxic effects from the metabolites

being accumulated. Several genes from some medicinal plants have been silenced using this technique such as *Papaver somniferum* (Hileman *et al*., 2005; Desgagné-Penix and Facchini, 2012; Lee and Facchini, 2010; Wijekoon and Facchini, 2012), *Eschscholzia californica* (Wege *et al*., 2007), *Aquilegia* (Gould and Kramer, 2007), and *Thalictrum* (Di Stilio *et al*., 2010). VIGS was also used to suppress the CYP80F1 family that may convert littorine to hyoscyamine in *Hyoscyamus niger*, tropane alkaloid producing species (Li *et al*., 2006). The method using pTRV vector system for silencing in *C. roseus* has been developed when three known steps in vindoline biosynthesis are silenced (Liscombe and O'Connor, 2011). The *N*-methyltransferase in *C. roseus* was successfully silenced and resulted in the accumulation of the NMT substrate of 16-methoxy-2,3-dihydro-3-hydroxytabersonine, and a decrease level of vindoline (Liscombe and O'Connor, 2011). An alternative method for pTRV vector delivery into *Catharanthus* plants has also been reported (De Luca *et al*., 2012b). Furthermore, the effectiveness of VIGS in silencing the target genes can be observed by suppressing the genes such as phytoene desaturase (PDS) that is involved in chlorophyll biosynthesis. The silencing of this gene produces visible white sectors on the affected leaf and serves as a marker that VIGS works effectively in the species being examined (Figure 1.5) (De Luca *et al*., 2012b).

VIGS technology has been applied for functional characterization of genes involved in *O*-demethylation of thebaine to codeine and of codeine to morphine in benzylisoquinoline alkaloid (BIA) biosynthesis of opium poppy (Hagel and Facchini, 2010). As VIGS experiments result in the accumulation of the substrates when the target gene is silenced, VIGS may provide the substrates required for the enzyme assays in

order to confirm the gene function by expressing the proteins in recombinant systems. In the case of silencing this opium poppy demethylase, the accumulated BIA intermediates can be purified by thin-layer chromatography and the substrates can be tested using recombinant proteins. Hence, VIGS offers another advantage by providing difficult-to-synthesize or commercially unavailable substrates for confirming the biochemical function of the candidate gene.

Figure 1.5 Virus-induced gene silencing of phytoene desaturase (PDS) in *Catharanthus roseus* (A) Seedlings treated with *Agrobacterium tumefaciens* strain GV3101 containing pTRV1 and pTRV2 vector control (B) Seedlings treated with *A. tumefaciens* containing pTRV1 and pTRV2 vector containing the PDS gene (GenBank™ accession: # JQ655739) fragment produce a photobleaching phenotype in leaves a few weeks after treatment (De Luca *et al*., 2012b).

VIGS is an efficient tool for gene function analysis in terms of narrowing down the large number of candidate genes that may be involved in the MIA biosynthesis pathway before performing detailed studies of enzyme characterization. As the number of available candidate genes continues to increase with the rise of the large-scale genomics technologies, VIGS offers a relatively rapid and efficient first screening step for MIA pathway gene identification. With the availability of this screening technique, *C. roseus* should continue to be a leading model system for studying MIA biosynthesis. However,

VIGS remains restricted to certain species and its efficiency has been mostly shown in the characterization of biosynthetic pathway genes, while genomics approaches lead to the identification of candidate genes from certain steps of the reactions in the pathway.

1.6 Metabolic engineering of MIA biosynthesis pathway

Identification of biosynthetic genes in the MIA biosynthesis pathway has permitted various efforts towards metabolic engineering. The goal of this type of technology is primarily to enhance the synthetic capacity of preferred products by the over-expression of pathway genes or by diverting pathways for the production of desired novel metabolites. This attempt requires insights into each reaction pathway step, its organization and its regulation. Some complications include those imposed by enzymes that may be key rate-limiting steps in the biosynthetic pathway. Proper compartmentalization may also be required to maintain suitable trafficking patterns. The first trial at metabolic engineering in MIA biosynthesis was the expression of constitutive TDC and STR transgenes using *C. roseus* cell cultures (Canel *et al*., 1998; Facchini, 2001). While over-expression of enzymes involved in the formation of tryptophan indole moiety, anthranilate synthase caused increase in the level of tryptophan without enhanced production of MIAs (Hong *et al*., 2006; Hughes *et al*., 2004), the monoterpenoid pathway has been considered the rate-limiting steps for MIA production. When the MEP pathway gene, 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) was over-expressed, alkaloid accumulated significantly, especially with one line that accumulated serpentine up to sevenfold compared to controls while no effect on catharanthine was observed (Ayora-Talavera *et al*., 2002). Moreover, studies with hairy root cultures over-expressing DAT

caused accumulation of hörhammericine compared with control roots (Magnotta *et al*., 2007). This study shows that the potential interactions between DAT and the root-specific MAT might affect the conversion of hörhammericine to 19-*O*-acetyl-hörhammericine (Figure 1.1). In terms of regulatory control of the MIA pathway, ORCA3 was suggested to activate the expression of anthranilate synthase, TDC, STR, D4H, cytochrome P450 reductase (CPR), and *D*-1-deoxyxylulose 5-phosphate synthase, but not G10H, SGD, and DAT (van der Fits and Memelink, 2000). Interestingly, in this constitutive expression of ORCA3 system, addition of exogenous loganin was required to increase MIA levels. Therefore, coordinated over-expression of transcriptional activators, ORCA2 and ORCA3 with the silencing of repressors like G-box binding factors 1 and 2 and the zinc finger protein family might be a solution to activate the MIA pathway (Menke *et al*., 1999; Pauw *et al*., 2004; Siberil *et al*., 2001; van der Fits and Memelink, 2000). Although various approaches of metabolic engineering in cell culture system have been reported, whole-plant metabolic engineering remains interesting for improving the productivity of MIAs, especially the potential of mutation breeding by screening for *C. roseus* germplasm that shows interesting genotypes, such as high level of vinblastine and low vindoline production. In this case, a low vindoline line (Magnotta *et al*., 2006) can be useful for further investigation of the later MIA biosynthesis pathway.

Once the genes in certain parts of a pathway are characterized, these multiple steps can be introduced in a suitable system such as yeast to improve the productivity of desired metabolites. Alternatively plants can be engineered to produce non-natural MIAs that can be useful to produce new drugs with MIA backbones while reducing the side effects of the drugs. For example, TDC expressed in transformed roots with chlorinated

tryptophan was able to produce chlorotryptamine derivatives and eventually chlorinated MIAs such as 10-chlorocatharanthine and 15-chlorotabersonine (Runguphan, Qu, and O'Connor, 2010; O'Connor, 2012). This example suggested that ability of plant to serve as factories for producing useful or even rare alkaloids, although their efficiency in plant systems needs to be further tested.

Several biosynthetic as well as regulatory genes have been used to genetically modify the production of MIAs in *C. roseus*. However, the most stable transformation methods in this species have been developed for roots and cell cultures (Pasquali, Porto, and Fett-Neto, 2006; Zárate and Verpoorte, 2007; Zhao and Verpoorte, 2007). While some success in plant cell culture technology for large-scale production of secondary metabolites has been reported, some drawbacks associated with this effort are low or unstable production of certain targeted metabolites, for example, the incapabilities of *C. roseus* cell culture to produce vindoline (Zhao and Verpoorte, 2007).

Another type of metabolic engineering study involves the substrate specificity of the enzyme (Ziegler and Facchini, 2008). Some studies that focused on the enzyme molecular structures use the crystal structure to generate the rational site-directed mutation. Mutant STR can also be coupled to SGD and resulted in a large alkaloid library that can be screened for new drugs. For instance, the crystal structure of STR complexed with strictosidine was analyzed to produce rational site-directed mutations. Of several STR mutants tested for their substrate specificity, one showed better turnover of some secologanin analogs compared with wild type STR (Chen *et al*., 2006; McCoy and O'Connor, 2006). Remarkably, rational two site-directed STR mutants were able to generate several strictosidine analogs with various substitutions that include halogenated

derivatives (Bernhardt, McCoy, and O'Connor, 2007; Ma *et al*., 2006). These analogs could eventually be fed to *C. roseus* hairy root cultures to produce modified MIAs.

Considering that certain parts of MIA biosynthesis genes have been elucidated, the current efforts are focused on transferring known parts of pathways to organisms such as *Saccharomyces cerevisiae* (Facchini *et al*., 2012). Some examples of the ability of yeast to carry out multiple-step plant pathways have been reported (Facchini *et al*., 2012) such as the production of the antimalarial drug, artemisinic acid in engineered yeast (Ro *et al*., 2006). While the major advantages of yeast as a production system have been well advertised in the literature, the problems solved by initial efforts to produce artemisinic acid will help to facilitate expression of large and more complex pathways such those for MIAs. Metabolic engineering of MIA biosynthesis is likely to reveal new problems that will have to be addressed including the challenge of functional expression of dozens of genes within the heterologous system and the expression of steps that may be regulated by light. The importance of compartmentalization, metabolite trafficking, and regulation of the MIA pathway *in planta* also need to be considered for the successful integration of this pathway in yeast.

1.7 Conclusions and Perspectives

It is apparent from this review that MIA biosynthesis is highly organized. In *C. roseus*, the pathway takes place in at least three different cell types, namely the IPAP, the epidermis, and the laticifer/idioblast cells. This separation demonstrates that mobile metabolites are shuttled between cell types. However, the mechanism of the transport and what intermediates are translocated needs to be further investigated. So far, the major site

of MEP biosynthesis as well as the early steps of iridoid biosynthesis is within the IPAP cells. An unknown iridoid likely to be loganic acid is then transported to the epidermis where the two last steps of iridoid biosynthesis occur. While early MIA biosynthesis and the first two steps in vindoline biosynthesis are localized to the epidermis, elaboration of strictosidine aglycone to catharanthine or tabersonine need to be addressed to confirm that these steps occur within the epidermal cells. An unidentified intermediate in vindoline biosynthesis is then transported from the epidermal cells to the mesophyll cells and eventually to idioblast/laticifer cells which explains vindoline accumulation within the leaf. Furthermore, localization of catharanthine biosynthesis can be evidenced by the accumulation of this metabolite on the leaf surface. This spatial separation of catharanthine and vindoline explain the low-level production of dimeric MIA vinblastine. The different situation in root systems where the entire MIA pathway from early precursors to final products occurs in the same cell type needs to be more fully studied for comparative purposes. In addition, the MIA pathway can be induced by signaling molecules such as MeJA that leads to the regulation of pathway by ORCA3 transcription factors as it is evident that many genes in MEP, iridoid, and MIA pathways are up-regulated in response to MeJA.

Since the 1960s, the MIA biosynthesis pathways have been investigated by chemical approaches, including isotopic-labeling experiments. The target of this approach was mainly to determine the intermediates in the pathway. In the last few years, the majority of studies have been on developing tools for expanding the functional characterization of genes involved in the MIA biosynthesis, especially by producing large-scale transcriptomes of different MIA-producing plant species and improved

comparative bioinformatic approaches for identifying candidate genes. Although numerous MIA biosynthetic genes in *C. roseus* have been elucidated, the intermediate steps from strictosidine aglycone to ajmalicine, catharanthine and tabersonine remain uncharacterized at the biochemical and molecular level. The growing transcriptomic information from various MIA-producing species complemented by organ/tissue-specific transcriptomic data from certain species (such as the epidermal enriched cDNA library in *C. roseus*) is being used to create a list of candidate genes for functional characterization. The development of VIGS method has expedited the gene discovery process as this tool can be used to screen the candidate genes by analyzing the changes in MIA metabolite profiles. The accumulation of the MIA intermediates in VIGS-treated plants may also supply the potential substrates that can then be purified to complete the biochemical characterization of the candidate gene in recombinant systems such as microorganisms or transgenic plants.

Future research could include the investigation of spatial localization of the MIA pathway within other Apocynaceae members if there are similar patterns of compartmentalization as in *C. roseus* observed in other species. A detailed understanding of metabolic regulation that is also influenced by complex organization with intra- and inter-cellular translocation of intermediates is essential to achieve the goal of metabolic engineering of MIA metabolism that may involve the assembly of the entire pathway in microorganism for large-scale production of valuable MIAs. The expansion of our knowledge of genomic studies, biochemistry, and molecular biology of MIA biosynthesis will facilitate the discovery of novel genes that can be utilized in engineered system for manufacturing biologically active MIAs as well as promoting the drug discovery process.

1.8 Acknowledgements

This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (V.D.L.), NSERC/BARD/Agriculture Canada team grant, Canada Research Chairs (V.D.L.), Genome Canada, Genome Alberta, Genome Prairie, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners.

Chapter 2

Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle

This manuscript was published in *Proceedings of the National Academy of Sciences of the United States of America* (2010) 107, 15287-15292

Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle

Jonathon Roepke[a,1], Vonny Salim[a,1], Maggie Wu[a], Antje Thamm[a], Jun Murata[b], Kerstin Ploss[c], Wilhelm Boland[c], and Vincenzo De Luca[a]

[a]Department of Biological Sciences, Brock University, St. Catharines, ON, Canada L2S 3A1; [b]Suntory Institute for Bioorganic Research, Osaka 618 8503, Japan, and [c]Max-Planck-Institut für Chemische Ökologie, 07745 Jena, Germany

[1] J.R. and V.S. contributed equally to this work

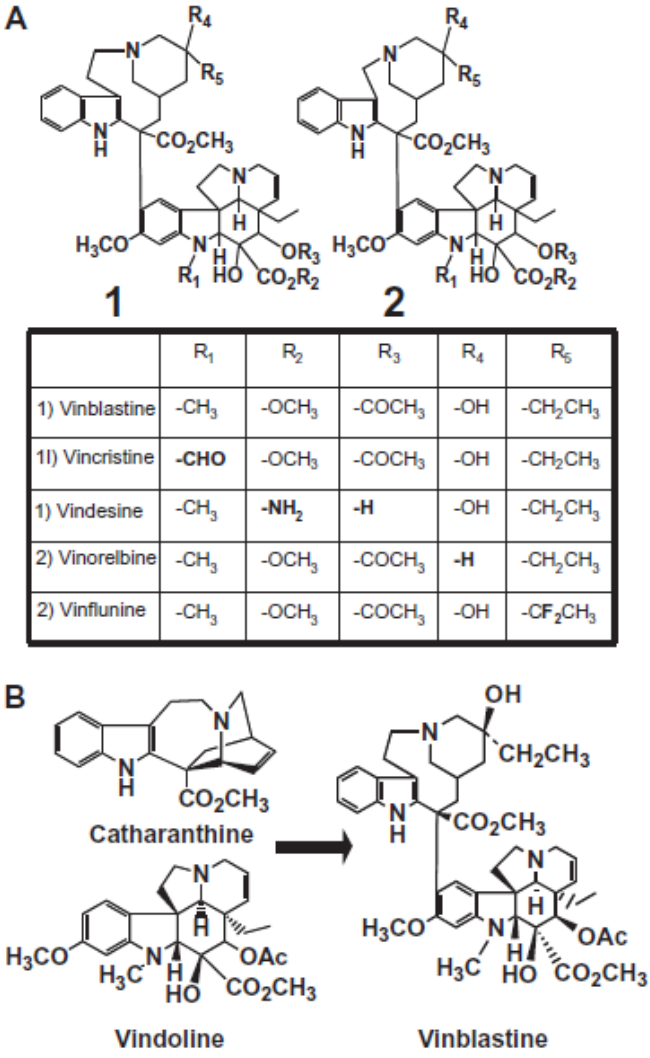
2.1 Abstract

The monoterpenoid indole alkaloids (MIAs) of Madagascar periwinkle (*Catharanthus roseus*) continue to be the most important source of natural drugs in chemotherapy treatments for a range of human cancers. These anticancer drugs are derived from the coupling of catharanthine and vindoline to yield powerful dimeric MIAs that prevent cell division. However, the precise mechanisms for their assembly within plants remain obscure. Here we report that the complex development-, environment-, organ-, and cell-specific controls involved in expression of MIA pathways are coupled to secretory mechanisms that keep catharanthine and vindoline separated from each other in living plants. Although the entire production of catharanthine and vindoline occurs in young developing leaves, catharanthine accumulates in leaf wax exudates of leaves, whereas vindoline is found within leaf cells. The spatial separation of these two MIAs

provides a biological explanation for the low levels of dimeric anticancer drugs found in the plant that result in their high cost of commercial production. The ability of catharanthine to inhibit growth of fungal zoospores at physiological concentrations found on the leaf surface of *Catharanthus* leaves, as well as its insect toxicity, provide an additional biological role for its secretion. We anticipate that this discovery will trigger a broad search for plants that secrete alkaloids, the biological mechanisms involved in their secretion to the plant surface and the ecological roles played by them.

2.2 Introduction

The vinca alkaloids are a well known source of drugs derived from the Madagascar periwinkle (*Catharanthus roseus*). The four major vinca alkaloids used in various cancer chemotherapies are vinblastine, vincristine (or semisynthetic derivatives), vindesine, and vinorelbine (Figure 2.1A). In addition, clinical trials are continuing with new semisynthetic derivatives such as vinflunine (Figure 2.1A) that may have improved biological properties and may be effective in a broader range of cancers. Although remarkable efforts by chemists have yielded the total synthesis of dimeric alkaloids (Mangeney *et al*., 1979; Kutney *et al*., 1988; Kuehne, Matson, and Bornmann, 1991; Magnus *et al*., 1992; Yokoshima *et al*., 2002), they or their monomeric precursors, catharanthine and vindoline, are still harvested and purified from periwinkle plants (Figure 2.1B), and the process of alkaloid extraction and purification produces low yields at very high cost.



	R_1	R_2	R_3	R_4	R_5
1) Vinblastine	$-CH_3$	$-OCH_3$	$-COCH_3$	-OH	$-CH_2CH_3$
1I) Vincristine	**-CHO**	$-OCH_3$	$-COCH_3$	-OH	$-CH_2CH_3$
1) Vindesine	$-CH_3$	**$-NH_2$**	**-H**	-OH	$-CH_2CH_3$
2) Vinorelbine	$-CH_3$	$-OCH_3$	$-COCH_3$	**-H**	$-CH_2CH_3$
2) Vinflunine	$-CH_3$	$-OCH_3$	$-COCH_3$	-OH	$-CF_2CH_3$

Figure 2.1 Commercially relevant dimeric MIAs (A) are derived by further modification of a dimer produced from the chemical or enzymatic coupling of catharanthine and vindoline (B).

The complexity of the vinca alkaloids is also matched by the multistep pathway involved in their assembly. This pathway requires tryptamine and the monoterpene secologanin for alkaloid assembly, and the pathway is highly regulated by development-, environment-, organ-, and cell-specific controls (Facchini and De Luca, 2008) that are poorly understood. Although biosynthesis of complex alkaloids such as tabersonine and catharanthine seem to occur within the protoderm and cortical cells of *Catharanthus* root

tips (Laflamme, St-Pierre, and De Luca, 2001), the pathway appears to be compartmented in multiple cell types in above-ground organs such as stems, leaves, and flowers (St-Pierre, Vazquez-Flota, and De Luca, 1999). The 2-C-methyl-D-erythritol-4-phosphate pathway, which supplies the carbon skeletons for secologanin biosynthesis (Contin *et al*., 1998) as well as geraniol 10-hydroxylase (G10H) involved in the first committed step in this pathway, have been localized to biochemically specialized internal phloem-associated parenchyma (IPAP) cells (Burlat *et al*., 2004; Mahroug, Burlat, and St-Pierre, 2007; Oudin *et al*., 2007b). In contrast, loganic acid *O*-methyltransferase (LAMT) and secologanin synthase that catalyze the terminal reactions in secologanin biosynthesis are expressed exclusively within the epidermis of young leaves and stems (Murata *et al*., 2008). This suggests very strongly that an uncharacterized pathway intermediate is transported between IPAP and epidermal cells to elaborate the secologanin molecule. The epidermis of leaves, stems, and flower buds also express tryptophan decarboxylase (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata and De Luca, 2005), strictosidine synthase (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata and De Luca, 2005), strictosidine β-glucosidase (Murata and De Luca, 2005), tabersonine 16-hydroxylase (Murata and De Luca, 2005), and 16-hydroxytabersonine 16-*O*-methyltransferase (16-OMT) (Murata and De Luca, 2005; Levac *et al*., 2008), whereas the *N*-methyltransferase, dioxygenase, and acetyltransferase (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata and De Luca, 2005) responsible for the last three steps in vindoline biosynthesis are expressed within leaf mesophyll cells or in specialized idioblasts and laticifers.

Carborundum abrasion has been used as a unique and complementary approach to obtain epidermis-enriched leaf extracts to measure alkaloid metabolites, enzyme activity,

and gene expression levels (Murata and De Luca, 2005; Levac *et al*., 2008). RNA analysis through random sequencing of the leaf epidermis has shown that this single layer of cells is biochemically rich in a variety of biosynthetic pathways that include those for flavonoid, very long chain fatty acids, pentacyclic triterpenes, and monoterpenoid indole alkaloids (MIAs), respectively (Murata *et al*., 2008). This biosynthetic diversity in a single cell type coincides with uncharacterized mechanisms that deliver lipids and triterpenes to the leaf surface to form the cuticle that seals the plant surface (Samuels, Kunst, and Jetter, 2008) and advanced alkaloid precursors to specialized leaf mesophyll idioblasts and laticifers (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata *et al*., 2008; Murata and De Luca, 2005) for elaboration into vindoline. The present study shows that directional transport mechanisms are coupled to alkaloid biosynthesis in the leaf epidermis to allow secretion of catharanthine into the surface wax while vindoline accumulates within specialized idioblast/ laticifer cells in the leaf. This spatial separation of MIAs provides a clear explanation for the low levels of dimeric anticancer alkaloids found in *Catharanthus* plants.

2.3 Methods

2.3.1 Leaf sample preparation

C. roseus plants were grown in a greenhouse under a 16/8-h light/dark photoperiod at 30 °C. Fresh young leaves from the third developmental stage were harvested and their fresh weights were recorded. To obtain surface extracts containing waxes, triterpenes and MIAs, leaves were dipped in 5 mL chloroform (Gniwotta *et al*., 2005) in 15-mL conical sterile capped polypropylene tubes (Sarstedt) for different times

(30 s to 1 h) at room temperature. The stripped leaves were removed and surface extract samples were evaporated to dryness by vacuum centrifugation in an SPD SpeedVac (Fisher Scientific). Before analysis by UPLC-MS, the dry materials were resuspended in 5 mL methanol. Stripped leaves were air-dried in a fume hood for 60 min to remove chloroform residues before they were extracted for MIAs.

2.3.2 Extraction of MIAs from leaves

Intact leaves or chloroform surface stripped leaves were frozen in liquid nitrogen, pulverized with a mortar and pestle, and homogenized with 5 mL methanol or chloroform. As catharanthine and vindoline from stripped or intact leaves were equally well extracted in either solvent system, all extractions were performed with methanol. The extracts were shaken at 100 rpm using an Innova 2000 shaker (New Brunswick Scientific) at room temperature for 1 h. The samples were mixed by vortex (Genie 2 vortex set at 10; Fisher Scientific) and centrifuged at 5,000 × g for 10 min to separate leaf debris and methanol. The methanol extracts were harvested and stored at −20 °C until further analysis. An aliquot of 200 μL of each extract was filtered through (0.22 μm) PALL filter (VWR) before analysis by UPLC/single-quadrupole MS (Waters).

2.3.3 UPLC-MS analysis of MIAs

The analytes were separated using an Aquity UPLC BEH C_{18} column with a particle size of 1.7 μm and column dimensions of 1.0 ×50 mm. Samples were maintained at 4 °C and 5-μL injections were made into the column. The analytes were detected by photodiode array and MS. The solvent systems for alkaloid analysis were as follows:

solvent A, methanol: acetonitrile:5-mM ammonium acetate and 6:14:80; solvent B, methanol: acetonitrile:5-mM ammonium acetate at 25:65:10. The following linear elution gradient was used: 0–0.5 min 99% A, 1% B at 0.3 mL/min; 0.5–0.6 min 99% A, 1% B at 0.4 mL/min; 0.6–7.0 min 1% A, 99% B at 0.4 mL/min; 7.0–8.0 min 1% A, 99% B at 0.4 mL/min; 8.0–8.3 min 99% A, 1% B at 0.4 mL/ min; 8.3–8.5 min 99% A, 1% B at 0.3 mL/min; and 8.5–10.0 min 99% A, 1% B at 0.3 mL/min. Alkaloid reference standards were also analyzed by using this method and to establish standard curves to quantify the levels of catharanthine, vindoline, and anhydrovinblastine in the extracts.

The mass spectrometer was operated with a capillary voltage of 2.5 kV, cone voltage of 34 V, cone gas flow of 2 L/h, desolvation gas flow of 460 L/h, desolvation temperature of 400 °C, and a source temperature of 150 °C. Catharanthine was detected at 280 nm and its identity was verified by its diode array profile, its mass (337 *m/z*), and retention time (4.45 min) compared with authentic standard. Vindoline was detected at 305 nm and its identity was verified by its diode array profile, its mass (457 m/z) and retention time (4.55 min). The 3′,4′-anhydrovinblastine was detected at 270 nm and its identity was verified by its diode array profile, its mass (793 m/z), and its retention time (5.7 min). Chromatographic peaks were integrated and compared with standard curves for catharanthine, vindoline and 3',4'- anhydrovinblastine to give the total amount of alkaloids in each sample.

2.3.4 HPLC-MS analysis of triterpenes

Triterpenes were extracted and analyzed by HPLC-MS as described previously (Murata *et al.*, 2008). Ursolic acid standard (Mr 456) (Sigma-Aldrich, Toronto, ON,

Canada) and *C. roseus* triterpenoids (crude extracts or samples partially purified by thin layer chromatography (TLC) were submitted to analysis using a Bruker HCT+LC/MS System (Mode:ESI-Negative Inlet Agilent 1100 LC; Solvent System: Acetonitrile/Water; Column: 50 mm C_{18} with a pre-column guard). Both crude and partially TLC purified triterpenoid samples displayed similar retention times (RT 16.0-16.2 and 16.4-16.6 min) and identical masses (455.3 *m/z*) as authentic ursolic acid standard.

2.3.5 Chlorophyll determination

For total chlorophyll content determination, 200 µL of each extract was mixed with 1.8 mL acetone, incubated for 30 min at −20 °C, and centrifuged for 20 min at 21,000 × g. The absorbance A645/A663 of the supernatant was determined by UV/VIS spectrophotometry zeroed at 750 nm (Ultrospec 3100 pro; Fisher Scientific).

2.3.6 Extraction of whole leaves for enzyme assays

Fresh young leaves from greenhouse-cultivated *C. roseus* were harvested (6.0 g) and frozen in liquid nitrogen. The leaf material was pulverized and homogenized in 12 mL of extraction buffer [100 mM Tris-HCl (Sigma), pH 7.5, 14 mM β-mercaptoethanol (EMD), and 25 mM KCl (Caledon)]. The homogenate was filtered through two layers of nylon and centrifuged at 500 × g (21000R; Fisher Scientific) for 10 min; the supernatant was added directly to a PD-10 column (GE Healthcare) for desalting. The desalted protein was used for enzyme assays and the protein concentration was determined with Bradford assays.

2.3.7 Extraction of epidermis-enriched proteins for enzyme assays

Fresh young leaves from greenhouse-cultivated *C. roseus* were harvested (6.0 g). The leaf material was added batch-wise to 60 mL of extraction buffer [100 mM Tris-HCl (Sigma), pH 7.5, 14 mM β-mercaptoethanol (EMD), and 25 mM KCl (Caledon)] containing 6 g of carborundum. The leaves, carborundum, and buffer were divided equally into two 50-mL tubes (Sarstedt) and vortexed (Genie 2; Fisher Scientific) at speed level 4 for 3 min. The leaf material was removed from the buffer and the samples were then centrifuged at 4,500 × g (21000R; Fisher Scientific) to separate the carborundum. The buffer (25 mL each) was then filtered via vacuum filtration through a 0.45-μm PALL membrane. The filtered material was then concentrated (Millipore) to a final volume of 5 mL. The concentrate was then desalted on a PD-10 column (GE Healthcare) and was used directly for enzyme assays and for protein determination with Bradford assays.

2.3.8 Geraniol 10-hydroxylase (G10H) assay

Enzyme assays contained 1 mL of epidermis-enriched or whole leaf extracts, geraniol (1 mM; Sigma), and NADPH (1 mM; Sigma), and further aliquots of NADPH (1 mM) were added at 10-min intervals. Assays were incubated at 37 °C (water bath, IsoTemp; Fisher Scientific) for 60 min. Additional control assays were performed without geraniol, without NADPH, or with denatured boiled extracts.

After terminating enzyme assays by addition of 500 μL ethyl acetate (Caledon), they were vortexed (Genie 2; Fisher Scientific) at speed level 10 for 1 min and the organic/aqueous phases were separated by centrifuged at 10,000 × g (Micromax RF;

Fisher Scientific). The top organic layer was harvested, and after repeating the extraction with another 500 μL ethyl acetate, the organic layers were pooled, evaporated to dryness (Speed Vacuum, Thermo Savant; Fisher Scientific) and resuspended in 200 μL methanol and analyzed by reversed-phase HPLC (2,996-photodiode array detector; Waters) using an Inertsil ODS-3 C_{18} column (4 × 250mm; GL Sciences) equipped with a 3 × 4-mm guard column (Phenomenex). Reaction products were injected onto the column and were eluted (1 mL min^{-1}) with solvent A (MeOH:MeCN:5 mM ammonium acetate at 6:14:80) and solvent B (MeCN; 100) using the following linear gradient: 0–20 min 80–20%A/20–80%B; 20–25 min 20–80% A/80–20% B; 25–35min 80–80% A/ 20–20% B. Geraniol and 10- hydroxygeraniol were monitored by photodiode array detection at 210 nm.

2.3.9 LAMT, 16-OMT, NMT, and DAT enzyme assays

LAMT, 16-OMT, and NMT radioactive assay measurements were performed as described previously (Murata *et al*., 2008). Enzyme assays (120 μL) contained LAMT (1.3 mM; Chromadex), 16-hydroxytabersonine (60 μM) or 2,3- dihydro-3-hydroxy-N-methyltabersonine (60 μM), 3.7 × 10^{-3} kBq [$^{14}CH_3$]-S-adenosyl-L-methionine (2.00 GBq/mmol; GE Healthcare), 100 μL of epidermis-enriched or whole leaf extracts, and were incubated at 37 °C for 60 min. DAT assays contained (120 μL) deacetylvindoline (60 μM), 3.7 × 10^{-3} kBq [^{14}-C]-acetyl-CoA (2.18 GBq/mmol; GE Healthcare), 100 μL of epidermis-enriched or whole leaf extracts and were incubated at 37 °C for 60 min. After LAMT assays were stopped by quick-freezing the assays in liquid nitrogen, they were lyophilized and dissolved in methanol (15 μL). Samples were submitted to silica gel TLC (Polygram Sil G/UV254; Macherey-Nagel) and radiolabeled loganin was resolved in (7:3

chloroform:methanol). Radioactive products were stored on a phosphor screen (GE Healthcare) for 16 h and emissions were detected using a phosphorimager (FLA-3000 equipped with multigauge version 3.0 software; Fujifilm). After harvesting the radioactive loganin spot, its concentration was measured by liquid scintillation counting (Beckman Coulter). Enzyme assays for 16-OMT, NMT, and DAT were stopped with 20 µL of 10M NaOH and alkaloids were extracted and processed as described for the G10H assay to produce evaporated residues. Reaction products from each assay were dissolved in methanol (10 µL). Samples were submitted to silica gel TLC (Polygram Sil G/UV254; Macherey- Nagel) and reaction products were resolved in ethyl acetate: methanol at a 9:1 ratio. The radioactive products were detected and their concentrations were measured as described for LAMT.

2.3.10 Effect of catharanthine on growth of Phytophthora nicotianae zoospores

P. nicotianae zoospores were prepared by aseptically removing a mycelia plug from a stock plate to fresh V8 agar medium. This medium was prepared by mixing well 10% vol/vol V8 juice, 1 g/L calcium carbonate, and 1.45% wt/vol agar, followed by autoclaving at 121 °C for 20 min, cooling to 40 °C, and pouring into sterile Petri plates that were stored at 4 °C until required. Mycelia were cultivated under continuous light for 7 d at 26 °C (TC19; Enconair Ecological Chamber). The culture was then treated with 0.5 to 1.0 mL of 0.01 M KNO_3 and incubated for a further 8 d under the same conditions. The plate containing mycelia was rinsed with sterile distilled water (10 mL) and air bubbles on the agar surface were dislodged by gently scraping the surface with a sterile plastic rod. After discarding the water, the washing procedure was repeated a second time.

Fresh sterile distilled water was added and the plate was transferred to continuous light for at least 2 h at room temperature. The plate was chilled on ice for 30 min and then placed at room temperature for a further 30 min to release zoospores. The liquid containing zoospores was transferred to a sterile test tube to be shaken vigorously on a vortex (Genie 2; Fischer Scientific) at speed level 5 for 2 min and then allowed to settle by gravity for 3 to 5 min. The concentration of zoospores in the supernatant was determined by counting them under a light microscope using a hemacytometer. Zoospores of a given concentration were transferred to 1.5 mL Eppendorf tubes containing 1 mL of half strength Gamborg B5 media (Gamborg *et al*., 1976), sucrose (20 g/L), myo-inositol (50 mg/L), nicotinic acid (0.5 mg/L), pyridoxine (0.5 mg/L), thiamine (5 mg/L; Sigma), and various concentrations of catharanthine to be cultivated under continuous light at 26 °C. The growth and multiplication of zoospore was determined in triplicate on days 0, 2, and 4 after the beginning of the experiment.

2.3.11 Effect of catharanthine on growth of insect larvae

Cultivation of insect larvae

The larvae of *Spodoptera littoralis* Boisd (Lepidoptera, Noctuidae) and *Spodoptera eridania* were reared with a 14 to 16 h photoperiod at 22 °C to 24 °C in plastic boxes using an artificial diet [300 g/L agar, 400 g/L bean flour, 3 g sodium ascorbate, 3 g ethyl p-hydroxybenzoate, 1 g formaldehyde (Bergomaz and Boppré, 1986), and added with 0.3 g β-sitosterol, 0.3 g leucine, 3.3 g vitamin mixture]. The larvae of *Helicoverpa armigera* were reared on artificial diet containing wheat germ and soybean powder between 24 °C and 27 °C with a 16:8 (L:D) photoperiod. The leaf bettles,

Phaedon cochleariae, were reared in the laboratory at 18 °C to 20 °C and 16:8 L:D period, fed on *Brassica rapa* spp. Chinensis (Kunert *et al*., 2008). The fifth-instar larvae of *Bombyx mori* (Lepidoptera) were reared at 14 to 16 h photophase between 22 °C and 24 °C, kept in plastic boxes, grown on artificial diet (mulberry leaf powder, potato starch, soybean powder, sucrose, agar, mineral salts, and vitamin B complex).

2.3.12 Feeding of Catharanthus leaves to insects

Four larvae of each species were placed in a closed Petri dish containing one *C. roseus* leaf pair in third developmental stage attached to its stem that was placed and sealed in an Eppendorf tube containing 1 mL of water.

Feeding of Mulberry artificial diets to Bombyx mori

C. roseus cv. Little Delicata leaves (2/3, 1, or 2 leaves) from the third developmental stage were frozen in liquid nitrogen and ground into a fine powder, then mixed with 2.8 g of Mulberry artificial diet that was used in feeding studies. Leaves from the same developmental stage were dipped in chloroform to obtain catharanthine-enriched extracts from the surface and the remaining leaf was extracted in methanol to obtain MIAs and other metabolites occurring inside the leaf. After evaporating both extracts to dryness, the MIAs were dissolved in 2.8 g of Mulberry artificial diet. In additional experiments, Mulberry artificial diet was supplemented in a similar manner with various concentrations of pure catharanthine for performing the feeding studies. Each of these diets was then analyzed for alkaloid content by analytical UPLC as described previously. Individual larvae were placed in a closed Petri dish and fed with a

fresh 0.2 g of refrigerated prepared diet containing different test components at 24-h intervals. The feeding study was continued for 7 d unless the larvae died before the end of the experiment. Each day, dead larvae were removed and stored in a 20 °C freezer for further dissection at the end of the experiment.

Collection of regurgitates

The regurgitates were collected and mixed with methanol at a 1:1 ratio after 2 and 16 h, centrifuged at 13,000 rpm for 15 min at 4 °C (Micromax RF; Thermo-IEC).

Collection of feces

The feces were collected after 30 h and pulverized with mortar and pestle in 1.5 mL of methanol.

Extraction of larval intestinal tracts, larval bodies, and remaining artificial diet mix

Larvae were dissected to separate their intestinal tracts from the rest of the body. Intestinal tracts were pulverized in 1 mL of methanol in 1.5 mL Eppendorf tubes and the supernatant was collected for analysis after centrifugation (Kunert *et al*., 2008). The larval body and remaining *Catharanthus* mixed artificial diet were homogenized individually using a mortar and pestle in the presence of 3 and 6 mL of methanol, respectively. The homogenate was transferred to 15-mL sealed plastic tubes and extracts were incubated on an Innova 2000 shaker (New Brunswick Scientific) at 125 rpm at room temperature for 1 h. An aliquot of 200 μL was then filtered through (0.22 μm) PALL filter (VWR) before being analyzed by UPLC/single-quadrupole MS (Waters).

2.4 Results

2.4.1 Waxy surface of C. roseus contains the complement of leaf catharanthine but not vindoline

Dipping of leaves in chloroform has been used to remove surface waxes from plant tissues (Samuels, Kunst, and Jetter, 2008) as well as other hydrophobic surface chemicals such as the pentacyclic triterpene ursolic acid that accumulates only on the surface of *Catharanthus* leaves (Murata *et al*., 2008). Analysis of chloroform soluble chemicals also showed that catharanthine occurs exclusively on the surface of these leaves (Figure 2.2A). Remarkably, only 2% to 5% of the vindoline could be extracted even if leaves were dipped in chloroform for as long as 1 h (Figure 2.2A). The identities and quantities of catharanthine and vindoline were verified by ultra-performance liquid chromatography (UPLC)–diode array detection (DAD)–MS (Figure 2.2B) compared with those of authentic catharanthine (*m/z* 337) and vindoline (*m/z* 457) standards (Figure 2.S1).

Analysis of the chloroform-soluble fraction showed that all the leaf catharanthine was on the leaf surface together with ursolic acid (Figure 2.2C), which is known to be present in the leaf cuticle (Murata *et al*., 2008), whereas only 3.8% of the vindoline appeared to be chloroform soluble. The extraction of chlorophyll and chlorogenic acid in chloroform was used to estimate if this solvent could extract metabolites from within leaf cells. Virtually no chlorophyll, a marker for leaf mesophyll metabolites, and less than 10% of the chlorogenic acid, a marker for leaf epidermis metabolites, were detected (Figure 2.2C). These data, together with the retention of a number of other metabolites

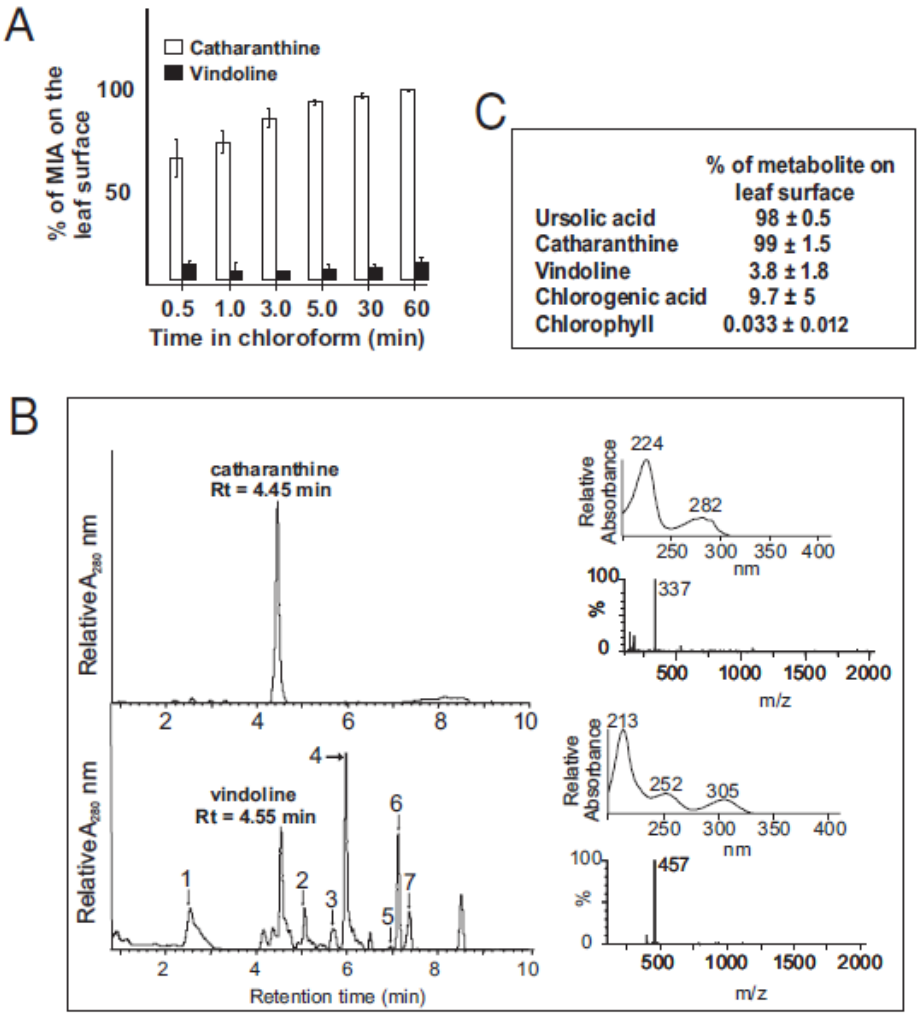


Figure 2.2 Catharanthine accumulates almost entirely in leaf wax exudates outside of the leaf epidermis, whereas vindoline is found within leaf cells. (A) Measurement of catharanthine and vindoline extracted from the leaf surface was made by dipping them in chloroform for different times. The experiment for each time point was performed in triplicate. The levels of catharanthine and vindoline per whole leaf representing the 100% value were 240 ± 9 μg and 113 ± 6 μg, respectively, in leaf number 3, as seen in Figure 2.3. (B) Chloroform-extracted MIAs and methanol-extracted MIAs remaining in leaves after chloroform treatments were measured by UPLC-MS (*Upper*). MIAs 1 to 7 were tentatively identified by their absorption and mass spectra as serpentine (1, *m/z* = 349); deacetoxyvindoline (2, *m/z* = 399); anhydrovinblastine (3, *m/z* = 794); unknown (4, *m/z* = 615); 16-methoxytabersonine (5, *m/z* = 367); unknown (6, *m/z* = 594); and unknown (7, *m/z* = 535 and 594). (C) Triplicate measurements of ursolic acid, catharanthine, vindoline, chlorogenic acid, and chlorophyll extracted from the leaf surface by dipping them in chloroform for 1 h. The levels per leaf pair of ursolic acid, catharanthine, vindoline, chlorogenic acid, and chlorophyll representing 100% were 919 ± 90, 271 ± 63, 120 ± 19, 210 ± 18, and 131 ± 17 μg/g fresh weight, respectively.

within the leaf (Figure 2.2B *Lower*), highlight the secretion of catharanthine to the leaf surface, whereas vindoline may be sequestered within idioblasts and laticifers (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata *et al*., 2008; Murata and De Luca, 2005).

The Mediterranean Deep orchid variety of *C. roseus* accumulates very low levels of vindoline because it lacks tabersonine 16-hydroxylase activity (Magnotta *et al*., 2006). The complement of catharanthine is also found on the surface of this variety when leaves were dipped in chloroform (Fig 2.S2A), whereas UPLC analysis of extracts of dipped leaf showed that only low levels of vindoline can be found within the leaf (Figure 2.2B and Figure 2.S2B). The decrease in vindoline content in this mutant line facilitated the detection of vindorosine by improving the absorption and mass spectra that could be obtained (Figure 2.S2C). These results confirm that, although a number of alkaloids accumulate within *Catharanthus* cells (Figure 2.2B and Figure 2.S2B) of these two cultivars, only catharanthine appears to be secreted to the surface.

2.4.2 Production of dimeric MIAs occurs only within older Catharanthus leaves

C. roseus (L.) G. Don cv. Little Delicata leaves of different ages were analyzed for their levels of catharanthine, vindoline, and 3', 4'-anhydrovinblastine (Figure 2.3). Although the concentrations of vindoline and catharanthine increased with leaf age from leaves 1 to 3, these MIAs leveled off and began to diminish from leaves 4 to 9. In all cases, virtually all the catharanthine was in the chloroform-soluble fraction whereas all the vindoline appeared only in extracts of whole leaves (Figure 2.3). The accumulation and secretion of catharanthine in the leaf wax exudate of young leaves strongly suggests that the pathway for catharanthine biosynthesis is also expressed in the leaf epidermis, as

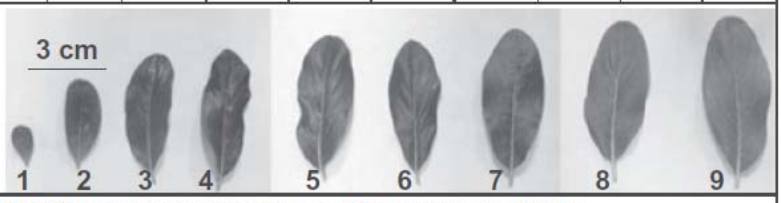
µg of MIA/leaf surface of Catharanthus roseus cv little delicata									
Leaf Surface	1	2	3	4	5	6	7	8	9
Catharanthine	23	122	246	194	175	151	139	125	93
Vindoline	0	0	0	3	0	0	0	0	0
3',4' Anhydrovinblastine	0	0	0	0	0	0	0	0	0
Leaf									
µg of MIA/leaf after extraction of leaf surface MIAs									
Leaf	1	2	3	4	5	6	7	8	9
Catharanthine	0	0	7	9	0	0	0	0	0
Vindoline	25	51	116	95	82	68	66	55	49
3',4' Anhydrovinblastine	0	0	16	18	25	19	0	0	0

Figure 2.3 Distribution of catharanthine and vindoline in the leaf wax exudate and within leaf cells in leaves of different ages

it is for 16-methoxytabersonine (Murata et al., 2008; Levac *et al*., 2008). Despite the fact that catharanthine and vindoline accumulated within separate locations, the levels of 3',4'-anhydrovinblastine did begin to increase within the fully expanded leaves 3 to 6 (Figure 2.3), whereas they disappeared in older leaves 7 to 9. The timing of 3',4'-anhydrovinblastine accumulation is consistent with earlier studies (Naaranlahti, Auriola, and Lapinjoki, 1991), suggesting that only older leaves were competent to produce or accumulate these dimers. It is interesting that small amounts of catharanthine were detected within leaves 3 and 4 (Figure 2.3), coinciding with the appearance of 3',4'-anhydrovinblastine in older leaves (Naaranlahti, Auriola, and Lapinjoki, 1991; Bowman, 1996). As the biosynthesis of MIAs, including catharanthine is restricted to younger leaves (Laflamme, St-Pierre, and De Luca, 2001; Murata *et al*., 2008) and to protoderm and cortical cells of root tips (St-Pierre, Vazquez-Flota, and De Luca, 1999), the source

of catharanthine inside older leaves is not known. It is possible that the catharanthine used for production of 3',4'-anhydrovinblastine is transported from root source, but the the transport of catharanthine from this underground source remains to be established. Alternatively, the internal source of catharanthine found in older leaves may be delivered by a switching in the transport of this molecule from the leaf epidermis toward the internal tissues of the leaf by transport mechanisms that remain to be determined.

2.4.3 Several Catharanthus species and hybrids secrete catharanthine to the leaf surface

The 100% surface accumulation of catharanthine was confirmed in leaves of several cultivars of *C. roseus* (L.) G. Don such as Little Delicata and Pacifica Peach, as well as in different species such as *Catharanthus longifolius* (Pichon) Pichon, *Catharanthus ovalis*, and *Catharanthus trichophyllus* (Baker) Pichon (Figure 2.4). These results, together with the idioblast/laticifer–specific expression of the last two steps in vindoline biosynthesis (St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata *et al*., 2008) and uncharacterized oriented transport mechanisms, provides a clear biological model for the sequestration of catharanthine and vindoline in separate leaf locations. As approximately 20% of plant species produce alkaloids, this result should trigger a broader search to discover if many more plants secrete alkaloids and to discover their biological/ecological roles on plant surfaces.

Madagascar periwinkle is also an extremely popular annual bedding plant because it blooms continuously throughout the growing season and there are many cultivars that produce a range of interesting colors. In addition, the plant is heat- and drought tolerant,

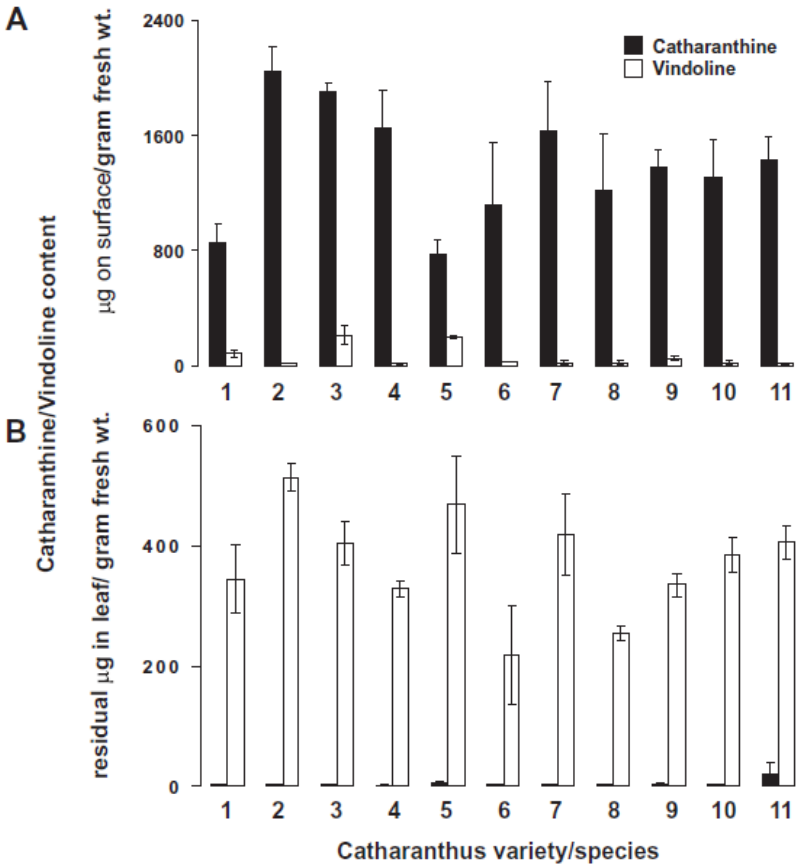


Figure 2.4 Distribution of catharanthine and vindoline on the surface and within the third leaf of various species/cultivars: *C. roseus* (L.) G. Don cv. Little Delicata (1), *C. roseus* Pacifica Peach (2), *Catharanthus longifolius* (Pichon) Pichon (3), *Catharanthus ovalis* (4), *Catharanthus trichophyllus* (Baker) Pichon (5), plus *C. roseus* hybrids (tolerant to the fungal pathogen *P. nicotianae* var. Parasitica) Cora Deep Lavender (6), Cora White (7), Cora Lavender (8), Cora Violet (9), Cora Apricot (10), and Cora Burgundy (11). The experiment for each time point was performed in triplicate.

which makes it a favorite for cultivation in the southern United States. Unfortunately, these cultivars are highly susceptible to fungal infections by *Phytophthora nicotianae* that causes aerial blight. This has led to the recent development of the Cora lines of *C. roseus* that appear to be resistant to this fungal pathogen. The mechanism of resistance in these new cultivars has been attributed to the production and accumulation of dimeric and trimeric MIAs (Bowman, 1996). All of the Cora cultivars analyzed (Figure 2.4; no. 6, deep lavender; no. 7, white; no. 8, lavender; no. 9, violet; no. 10, apricot; and no. 11,

burgundy) also accumulate all of their catharanthine on the leaf surface, whereas vindoline was almost exclusively found within leaves. Remarkably, the Cora cultivars accumulate similar levels of catharanthine and vindoline as in other *Catharanthus* cultivars (Figure 2.4) and detailed UPLC-MS analysis did not reveal the presence of trimeric alkaloids in the Cora cultivars.

2.4.4 Catharanthine has antifungal, antiinsect, and other biological properties

The spatial separation of vindoline and catharanthine appears to prevent dimer formation, perhaps because they may also interfere with cell division in *Catharanthus* as in cancer cells. The presence of catharanthine in the wax layer may also have a biological role, as shown with the MIA ibogaine that displays antifungal activity against *Candida albicans* (Yordanov *et al*., 2005). Incubation of catharanthine with zoospores of *P. nicotianae* inhibits both growth of zoospores and the formation of hyphae at a concentration of 10 µg/mL of culture medium (Table 2.1).

Table 2.1 Effect of catharanthine concentration on growth of *Phytophthora nicotianae*

	Day 0		Day 2		Day 4	
Catharanthine treatment, µg/mL	Zoospores, $\times 10^3$	Presence of hyphae	Zoospores, $\times 10^3$	Presence of hyphae	Zoospores, $\times 10^3$	Presence of hyphae
0	30	—	146 ± 6	+++	1,493 ± 9	++++
1*	30	—	520	++	400	++
10	30	—	20 ± 14	—	30 ± 14	—
50	30	—	20 ± 14	—	20 ± 14	—

The growth of *P. nicotianae* in the presence of different concentrations of catharanthine was monitored by counting the number of zoospores and by visual measurement of the abundance of hyphal growth after 2 and 4 d of cultivation compared with those growing in the presence of B5 medium.
*All experiments were performed in triplicate apart from those showing zoospore growth in 1 µg/mL of catharanthine, which was a single measurement.

The concentration of catharanthine/square centimeter of leaf surface in leaves 1 to 5 (Figure 2.3) varied between 23 μg in leaves 1 and 2 and 14 μg in leaf 5, sufficiently high to affect the activities of *P. nicotianae* zoospores on the leaf surface of *Catharanthus* species. In addition, the catharanthine but not the vindoline component of dimeric MIAs has been shown to interact directly with tubulin, and catharanthine alone can induce self-association of tubulin (Prakash and Timasheff, 1991). The binding constants to tubulin for vinblastine (4.3 μM), catharanthine (2.8 mM), and vindoline (50 mM) were described by Prakash and Timasheff (1991). However all the *C. roseus* lines that are resistant (Cora lines) or susceptible to this fungal pathogen contain similar levels of catharanthine (Figure 2.4). This suggests that, although catharanthine can inhibit the *in vitro* growth of *P. nicotianae* zoospores (Table 2.1), additional factors are involved in the *in vivo* disease resistance observed in Cora lines.

The potential biological role of catharanthine on herbivory was also tested by direct feeding of *Catharanthus* leaves to *Spodoptera littoralis*, *Spodoptora eridania*, *Helicoverpa armigera*, *Phaedon cochleariae*, and *Bombyx mori* (Table 2.S1). Although third instar larvae of *S. eridania* would not eat *Catharanthus* leaves, the two Spodoptora species, as well as *H. armigera*, were generally able to consume various amounts of these leaves without any insects dying during the course of the experiment. In contrast, both *P. cochleariae* and *B. mori* refused to eat *Catharanthus* leaves and died after 5 d without feeding.

It was found that sixth instar larvae of *B. mori* would consume pulverized *Catharanthus* leaf extracts when these were incorporated into commercially available Mulberry diet (Table 2.2). When the larvae were not fed anything, they all died within 2

d, whereas none died when they were fed their regular Mulberry diet over a 7-d period. In contrast, insects being fed Mulberry diet mixed with different amounts of *Catharanthus* leaves died progressively sooner in a dose-responsive manner (Table 2.2). Similarly, larvae died after 2 d when they were fed Mulberry diets containing catharanthine-enriched leaf surface extracts, whereas they died after 4 d of feeding with *Catharanthus* leaf extracts obtained after the catharanthine and other surface components were removed by chloroform treatment. The regurgitate, feces, intestinal tracts, and bodies of larvae as well as the remaining Mulberry/*Catharanthus* diet were collected to analyze the alkaloid content in each component (Table 2.S2).

Table 2.2 Effect of catharanthine concentration on fifth-instar *B. mori* larvae

Treatment*	Catharanthine, µg	Days before death
No food	0	2 ± 1
Mulberry diet	0	7
Mulberry diet + 2/3 catharanthus leaf	92.1 ± 8.6	5.3 ± 0.6
Mulberry diet + 1 catharanthus leaf	138.1 ± 12.8	4 ± 1
Mulberry diet + 2 catharanthus leaves	276.3 ± 25.7	3.3 ± 0.6
Mulberry diet + chloroform extracts of Leaf surface	180.5 ± 39.6	2 ± 0
Mulberry diet + extracts of chloroform treated leaf	1.4 ± 0.4	4.3 ± 0.6
Mulberry diet + catharanthine	50	4.7 ± 0.6
Mulberry diet + catharanthine	250	2.4 ± 1.5
Mulberry diet + catharanthine	500	1.7 ± 0.6

*Fifth-instar *B.mori* larvae were fed various diets in triplicate for a 1-wk period. Larvae not provided food died within 2 d of initiating the experiment compared with larvae fed a Mulberry diet that did not die during 7 d of observation.

Although no anhydrovinblastine could be found in the regurgitate, feces, intestinal tracts, and bodies of larvae, all contained detectable levels of catharanthine and vindoline, with the highest levels found in the larval body and in the feces. When larvae were fed Mulberry diets containing different amounts of catharanthine, they also died

progressively sooner in a dose-responsive manner (Table 2.2). These results suggest that the presence of catharanthine on the leaf surface could be an important deterrent to insect herbivory by causing decreased feeding or through its toxicity that appears to lead to insect death (Table 2.2).

2.5 Discussion

Plants are well known to secrete a range of secondary metabolites onto their surfaces that may offer a chemical barrier to insect herbivores or to attack by fungi and other pathogens. Although these leaf wax exudates may contain unusual fatty acids, triterpenes (Murata *et al*., 2008; Usia *et al*., 2005; Kubo and Matsumoto, 1984) and other terpenes, hydrophobic flavonoids (Valant-Vetschera and Brem, 2006) and other phenols, there is little evidence for the secretion of alkaloids onto the plant surface. Water-soluble pyrrolizidine alkaloids were recently reported on the surface of *Senecio jacobaea* (Vrieling and Derridj, 2003), although the levels found were small compared with those within the leaves. The present study shows that MIAs are clearly found on plant leaf surfaces, as the complement of catharanthine in several *Catharanthus* species accumulates in the leaf wax exudate (Figure 2.2, 2.3, and Figure 2.4). In contrast, vindoline, anhydrovinblastine, and the other MIAs accumulate only within leaf cells.

2.5.1 Secretion of catharanthine to the leaf surface expands the leaf cell types involved in the biosynthesis and accumulation of MIAs in Catharanthus

Previous studies have shown that silicon carbide abrasion techniques could be used to preferentially extract metabolites, RNA, or proteins for metabolomic,

transcriptomic, or enzymatic analysis (Murata and De Luca, 2005; Levac *et al*., 2008). This technique is used in the present study to determine the ratios of G10H; LAMT; 16-OMT; 16- methoxy-2,3-dihydro-3 hydroxy tabersonine-*N*-methyltransferase (NMT) and deacetylvindoline 4-*O*-acetyltransferase (DAT) enzyme activities in leaf epidermis enriched extracts compared with whole leaf extracts (Figure 2.5). As shown by previous studies (Murata and De Luca, 2005, Levac *et al*., 2005), the specific activities of LAMT and 16-OMT were more than 16- fold higher within leaf epidermis enriched extracts compared with whole leaves, whereas little or no G10H, NMT, or DAT could be detected within leaf epidermis enriched extracts (Figure 2.5). These results support the leaf epidermis as a major site of MIA biosynthesis.

The present studies have expanded the complexity of MIA biosynthesis in medicinal periwinkle and have provided insights into the leaf epidermal specialization of this pathway to assemble catharanthine and 16-methoxytabersonine (Figure 2.5) from the early precursors tryptamine and secologanin. The elaboration of these two metabolites, combined with the secretion of catharanthine into the wax layer (Figure 2.2, 2.3, 2.S2, and 2.4), the possible secretion of 16-methoxytabersonine or a later intermediate into the mesophyll layer, and leaf mesophyll-specific expression of the last three or four steps in vindoline biosynthesis (Facchini and De Luca, 2008; St-Pierre, Vazquez-Flota, and De Luca, 1999; Murata *et al*., 2008) (Figure 2.5), describes the biological mechanism that appears to be required for *Catharanthus* species to accumulate both end products in the same plant, but in two separate locations. The spatial separation of these two alkaloids also explains why plants do not accumulate higher levels of dimeric anticancer drugs such as anhydrovinblastine. The biological reasons for this may be related to the toxicity

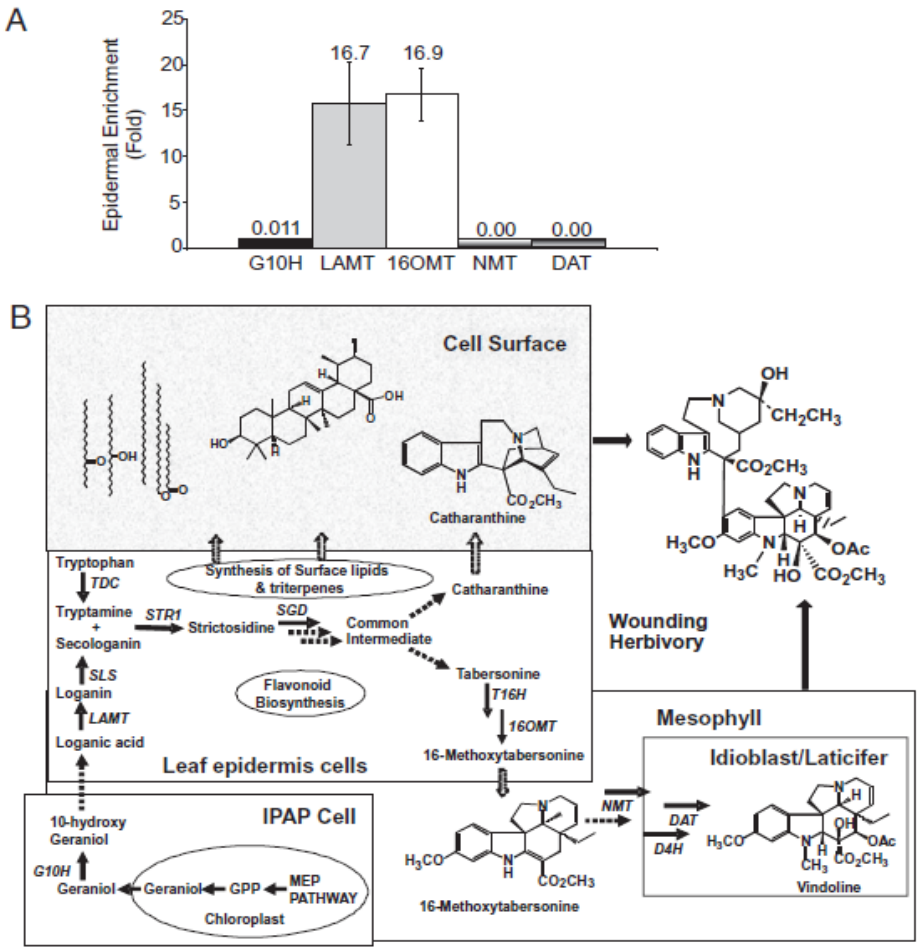


Figure 2.5 Model for biosynthesis and secretion of secondary metabolites produced in the epidermis of *C. roseus* leaves. (A) Enrichment of monoterpenoid indole alkaloid pathway enzyme activities in leaf epidermis enriched protein extracts produced by carborundum abrasion compared with those found in whole leaves. The whole leaf–specific activities used to establish the fold enrichment of enzymes in the leaf epidermis were a follows: G10H (184 ± 14 nmol/μg protein), LAMT (0.69 ± 0.02 pmol/μg protein), 16-OMT (1.89 ± 0.05 pmol/μg protein), NMT (0.47 ± 0.01 pmol/μg protein), and DAT (0.37 ± 0.02 pmol/ μg protein). (B) The leaf epidermis is specialized for biosynthesis of flavonoids, the fatty acid components of waxes, triterpenes, and MIAs. Although the 10-hydroxygeraniol required for biosynthesis of secologanin is made in specialized IPAP cells via the methyl–erythritol pathway (MEP) and G10H, this metabolite or another intermediate is transported to the leaf epidermis, where loganic acid is converted to secologanin via LAMT and secologanin synthase (SLS). Within the leaf epidermis, tryptophan is decarboxylated in the cytoplasm to tryptamine by tryptophan decarboxylase (TDC) and strictosidine is formed by the coupling of tryptamine and secologanin in the leaf epidermal vacuole by strictosidine synthase (STR1). Strictosidine is released into the cytoplasm as a reactive aglycone by cytosolic strictosidine β-glucosidase (SGD), where mostly unknown biochemical transformations lead to the production of the three major corynanthe, iboga, or aspidosperma—backbones of MIAs. A common intermediate is converted to catharanthine or to tabersonine by uncharacterized biochemical reactions. Virtually, all the catharanthine is secreted into the leaf cell surface, together with fatty acid components of waxes and the triterpene, ursolic acid. After tabersonine is converted into 16-methoxytabersonine in the leaf epidermis by tabersonine 16-hydroxylase (T16H) and by 2,3-dihydro-3-hydroxy-*N*-methyltabersonine (16OMT), this metabolite may be secreted into leaf mesophyll cells, where a further uncharacterized oxidation takes place, followed by a *N*-methylation (NMT) located in chloroplast thylakoids, hydroxylation (D4H), and *O*-acetylation (DAT) to yield vindoline. The last two reactions in vindoline biosynthesis appear to occur in specialized leaf idioblast and laticifer cells. It is postulated that wounding or herbivory may bring catharanthine and vindoline together to allow the formation of dimeric anticancer MIAs. The dotted lines represent uncharacterized reactions or unknown mechanisms that remain to be documented.

of anhydrovinblastine that may also interfere with cell division in *Catharanthus* as it does in cancer cells. Although this may not occur in fifth instar *B. mori* larvae (Table 2.2), it may be hypothesized that catharanthine and vindoline could be part of preformed chemical defenses that are mixed by wounding caused by herbivory (Figure 2.5), whereby they could combine in the intestinal tracts of the herbivore, where anhydrovinblastine would be produced in chemical or enzymatic dimerization. The mechanism of dimer formation in older *Catharanthus* leaves is obscure, but this may be associated with the appearance of the development-specific expression in older leaves of an alkaline class III peroxidase capable of converting catharanthine and vindoline into anhydrovinblastine (Costa *et al*., 2008). If such a mechanism is involved, a key question that remains is how catharanthine and vindoline could be placed in the same cellular compartment as this peroxidase. The appearance of low levels of catharanthine within older leaves (Figure 2.3) suggests this may be the case, but it is not clear from where the catharanthine is derived.

2.6 Conclusions

This study reveals that assembly of MIAs in *C. roseus* are remarkably dynamic, involving at least three cell types within leaves. Rates of biosynthesis, as well as oriented transport mechanisms, precisely regulate where, when, and how different MIAs accumulate during plant growth and development. The intimate association between cellular specialization and oriented secretion ensures that both catharanthine and vindoline are deposited into separate leaf locations, which explains the need for this unique and complex biological organization. Remarkably, *Catharanthus* leaves and/or

their leaf extracts deterred feeding by *Spodoptora littoralis* (Meisner et al., 1981) and *Spodoptera exigua* (Luijendijk, van der Meijden, and Verpoorte, 1996), whereas leaf extracts inhibited growth of *Spodoptera litura* (Deshpande, Joseph, and Sharma, 1988). The discovery of catharanthine on the leaf surfaces of four separate *Catharanthus* species suggests that many more plant species may actually secrete alkaloids for defensive reasons as well as other functions that remain to be discovered in nature.

2.7 Acknowledgements

This work was funded by a discovery grant from the National Sciences and Engineering Research Council of Canada and by a Tier 1 Canada Research Chair in Plant Biotechnology (to V.D.L.).

Chapter 3

Iridoid glucosyltransferase expressed in internal phloem-associated parenchyma cells is involved in the assembly of secologanin and monoterpenoid indole alkaloids in Madagascar periwinkle

This manuscript has been accepted for publication in *The Plant Cell*

Iridoid glucosyltransferase expressed in internal phloem-associated parenchyma cells is involved in the assembly of secologanin and monoterpenoid indole alkaloids in Madagascar periwinkle

Keisuke Asada[1,a], Vonny Salim[1,b], Sayaka Masada-Atsumi[1,a,b.c], Elizabeth Edmunds[b], Mai Nagatoshi[a], Kazuyoshi Terasaka[a] , Hajime Mizukami[a], and Vincenzo De Luca[b]

[a]Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya 467-863;

[b]Department of Biological Sciences, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, L2S 3A1, Canada;

[c]Division of Pharmacognosy, Phytochemistry and Narcotics, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo, 158-8501, Japan

[1]These authors contributed equally to the study

3.1 Abstract

Iridoids form a broad and versatile class of biologically active molecules found in thousands of plant species. In addition to the many hundreds of iridoids occurring in plants, some iridoids like secologanin serve as key building blocks in the biosynthesis of thousands of monoterpene indole alkaloids (MIAs) and many quinoline alkaloids. The present study describes the molecular cloning and functional characterization of three iridoid glucosyltransferases (CrUGT-6, -7, and -8) from Madagascar periwinkle (*Catharanthus roseus*) with remarkably different catalytic efficiencies. Biochemical

analyses reveal that CrUGT8 possessed a high catalytic efficiency toward its exclusive iridoid substrate, 7-deoxyloganetic acid, making it better suited for the biosynthesis of iridoids in *C. roseus* than the other 2 iridoid GTs. The role of CrUGT8 in the 4th to last step in secologanin biosynthesis was confirmed by virus-induced gene silencing in *Catharanthus* plants that reduced expression of this gene and resulted in a large decline in secologanin and MIA accumulation within silenced plants. Localization studies of *CrUGT8* by carborundum abrasion methods show that its expression occurs preferentially within *Catharanthus* leaves rather than in epidermal cells and *in situ* hybridization studies confirm that *CrUGT8* is preferentially expressed in internal phloem associated parenchyma cells of *Catharanthus roseus*.

3.2 Introduction

The Madagascar periwinkle (*Catharanthus roseus*) is biochemically specialized for production of numerous monoterpenoid indole alkaloids (MIAs) that are a feature of thousands of species from the Apocynaceae, Gentianaceae, Loganiaceae, and Rubiaceae plant families. The medical relevance of *Catharanthus* is attributed to the trace presence of the important anticancer drugs, vinblastine and vincristine that are derived from the coupling of the more abundant MIAs, vindoline and catharanthine found in above ground parts of the plant. These dimeric MIAs as well as a series of chemically modified derivatives act by disrupting cell division through their binding to microtubules and continue to be used today for the effective treatment of various cancers. Studies to understand why only low levels of dimeric MIAs are produced in *Catharanthus* plants have shown that catharanthine molecule is exported to the surface of leaves and is

spatially separated from vindoline that accumulates within specialized cells in the leaf mesophyll (Roepke *et al*., 2010).

Remarkably, the biosynthetic pathway for assembly of MIAs is also organized so that different cell types are specialized for various parts of the biosynthetic pathway. Biochemically specialized internal phloem-associated parenchyma cells (IPAP) preferentially express the methylerythritol phosphate pathway (Burlat *et al*., 2004) that provides geraniol through the action of an IPAP-associated geraniol synthase (Simkin *et al.*, 2013). Within these cells, geraniol is then converted to 10-hydroxygeraniol via the action of geraniol 10-hydroxylase (Burlat *et al*., 2004) (Figure 3.1) and after further oxidation to the dialdehyde, a novel iridoid synthase recruited from a short chain reductase gene family converts this intermediate into cis-trans-nepetalactol/iridodial (Geu-Flores *et al*., 2012). The remaining steps involve uncharacterized enzyme(s) (Figure 3.1D) that convert this intermediate into 7-deoxyloganetic acid (**3**, Figure 3.1), followed by glucosylation (Figure 3.1E) to form 7-deoxyloganic acid (**4**, Figure 3.1), hydroxylation (Figure 3.1F) and carboxy-*O*-methylation (Figure 3.1G) to form loganin (**6**, Figure 3.1). An alternative pathway has been proposed that involves carboxy-*O*-methylation (Figure 3.1J) of 7-deoxyloganetic acid (**3**, Figure 3.1) to form 7-deoxyloganetin (**8,** Figure 3.1) before glucosylation (Figure 3.1K) and hydroxylation (Figure 3.1L). A recent study that supports this latter pathway involved an iridoid glucosyltransferase (Nagatoshi *et al*., 2011) from *Gardenia jasminoides* that preferentially glucosylated the 1-*O*-hydroxyl group of 7-deoxyloganetin (**8**, Figure 3.1K) that had no activity towards 7-deoxyloganetic acid (**3**, Figure 3.1E). However, the

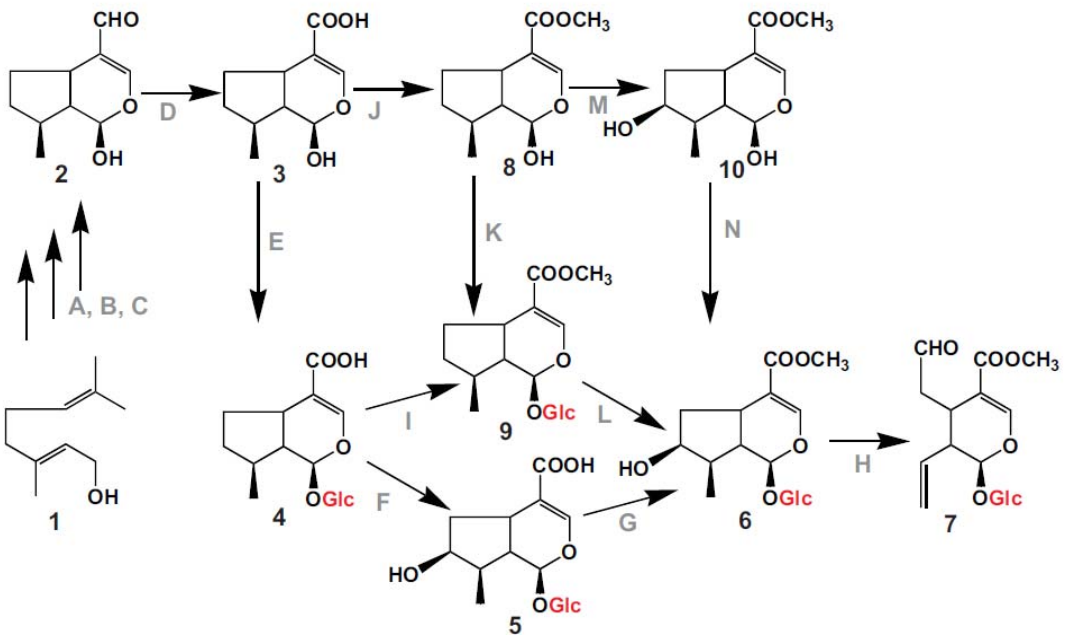


Figure 3.1 Possible secologanin biosynthesis pathways. Intermediates in secologanin biosynthesis pathway: **1)** geraniol, **2)** iridotrial, **3)** 7-deoxyloganetic acid, **4)** 7-deoxyloganic acid, **5)** loganic acid, **6)** loganin, **7)** secologanin, **8)** 7-deoxyloganetin, **9)** 7-deoxyloganin, **10)** loganetin. Enzymes involved in secologanin biosynthesis: **A)** *geraniol 10-hydroxylase (G10H/CYP76B6)*, **B)** *10-hydroxygeraniol oxidoreductase (10HGO)*, **C)** *iridoid synthase/monoterpene cyclase (IRS/MTC)*, **D)** iridodial oxidoreductase (IRDLOX), **E)** *7-deoxyloganetic acid glucosyltransferase (DLEGT)*, **F)** 7-deoxyloganic acid 7-hydroxylase (DL7H), **G)** *loganic acid methyltransferase (LAMT)*, **H)** *secologanin synthase (SLS/CYP72A1)*, **I)** 7-deoxyloganic acid methyltransferase (DLMT), **J)** 7-deoxyloganetic acid methyltransferase (DLEMT), **K)** 7-deoxyloganetin glucosyltransferase (DLENGT), **L)** 7-deoxyloganin 7-hydroxylase (DLN7H), **M)** 7-deoxyloganetin 7-hydroxylase (DLEN7H), **N)** loganetin glucosyltransferase (LENGT). The genes that have been cloned and functionally characterized are in italics and include CrUGT8 described in the present study.

preference of native and recombinant loganic acid *O*-methyltransferase (LAMT) from *C. roseus* (Murata *et al*., 2008) for loganic acid (**5**, Figure 3.1G), but not 7-deoxyloganic acid (**4**, Figure 3.1I) makes the alternative pathway less likely in *Catharanthus* plants. The last 2 steps to form secologanin involve LAMT and secologanin synthase (SLS) (Figure 3.1H) that are preferentially expressed in the leaf epidermis of *Catharanthus* plants as determined by carborundum abrasion (Murata *et al*., 2008) and *in situ* hybridization (Guirimand *et al*., 2011a) methods. The formation of secologanin in the

leaf epidermis suggests that an undefined iridoid, possibly loganic acid or some earlier intermediate must be transported to the leaf epidermis for its elaboration and this cell type is also the site of expression of enzymes such as tryptophan decarboxylase and strictosidine synthase that are required to elaborate the formation of strictosidine from which all of the MIAs of *Catharanthus* are derived (Facchini and De Luca, 2008; Guirimand *et al*., 2011a).

The glucosylation of a range of plant natural products are well known to be catalyzed by family 1 plant secondary product glycosyltransferases (PSPGs) defined by the presence of a 44 amino acid C-terminal motif known as a PSPG box (Vogt and Jones, 2000) that functions as a sugar donor binding pocket. Genipin glucosyltransferase (GjUGT2) from *Gardenia* was cloned by isolating a number of genes containing this conserved PSPG box and by its functional characterization in *E. coli* (Nagatoshi *et al*., 2011). By using the same approach, the present report further elucidates the pathway responsible for formation of secologanin by describing the isolation, biochemical and molecular characterization of 3 separate CrUGTs (CrUGT6, -7, and -8) that carry out iridoid glucosyltransferase reactions with remarkably different efficiencies. The study identifies the role of CrUGT8 in secologanin biosynthesis by showing that virus-induced gene silencing (VIGS) reduces expression of this gene and results in a large decline in secologanin and MIA accumulation within silenced plants. Localization studies of CrUGT8 by carborundum abrasion methods show that its expression occurs preferentially within *Catharanthus* leaves and *in situ* hybridization confirms that CrUGT8 is preferentially expressed in IPAP cells of *C. roseus* where iridoid biosynthesis is initiated.

3.3 Methods

3.3.1 Plant materials

Catharanthus roseus and *Catharanthus longifolius* were grown either in a greenhouse or in an incubator at about 25°C under continuous light. Cell suspension cultures of *C. roseus* were originally established from seedling-derived callus and maintained in in LS medium (Linsmaier and Skoog, 1965) supplemented with 3% sucrose, 1 μM 2,4-dichlorophenoxyacetic acid and 1 μM kinetin. The cells were cultured on a rotary shaker (100 rpm) at 25°C in the dark, and subcultured at two-week-intervals.

3.3.2 Chemicals

7-Deoxyloganin tetraacetate was kindly provided by Professor K. Inoue (Yokohama College of Pharmacy). 7-Deoxyloganin and 7-deoxyloganic acid were prepared from 7-deoxyloganin tetraacetate according to the method described previously (Nagatoshi *et al*., 2011). Loganetin, 7-deoxyloganetin, and 7-deoxyloganetic acid were enzymatically prepared from loganin, 7-deoxyloganin, and 7-deoxyloganic acid, respectively, as follows. A 20-mg aliquot of each glucoside was dissolved in 10 mL of 50 mM phosphate-citrate buffer (pH 4.8) containing 50 units/mL almond β-glucosidase (Sigma) and incubated for 3 h at 37°C. After centrifugation, the reaction mixture was extracted with ethyl acetate three times. The combined ethyl acetate extract was concentrated under reduced pressure to yield the aglycone. The purity of each aglycone obtained was estimated by quantitative ^{1}H NMR analysis (Hasada *et al*., 2011). Geniposide, genipin, and loganin were obtained from Wako Pure Chemicals (Osaka, Japan), and secologanin from Sigma. Iridotrial was synthesized by Dr. N. Kato, Nagoya

City University. All other chemicals were commercial reagent-grade quality unless otherwise stated.

3.3.3 Homology-based cloning of UGTs

Total RNA was isolated from *C. roseus* cultured cells or leaves using an RNeasy Plant Mini Kit (Qiagen). RT-PCR was performed using a CapFishing Full-Length cDNA Premix Kit (Seegene). Two degenerate primers UGT2mFw (5'-TTYBTIWSICAYTGYGGITGGAA-3') and PSPG2Fw (5'-TGYGGITGGAAYTCIRYIYTIGA -3') were designed based on the highly conserved amino acid sequences F(L/V)(T/S)HCGWN and CGWNS(T/V)LE, respectively, in PSPG box of plant glucosyltransferases (Vogt and Jones, 2000). A 5-μl aliquot of the cDNA was used as a template for PCR amplification in a 50-μl reaction mixture containing 1 μM primer UGT2mFw, 0.2 μM 3'-RACE primer from the CapFishing Full-Length cDNA Premix Kit, and 25-μl SeeAmp TaqPlus Master Mix (Seegene). A portion of the first PCR product was used as the template for nested PCR using the PSPG2Fw and the 3'-RACE primers. PCR was performed under the following conditions: denaturation at 94 °C for 3 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 1 min, and extension at 72 °C for 1 min, and final extension at 72 °C for 5 min. Amplified products of about 500 bp were recovered from agarose gel and subcloned into the pMD20 T-vector (Takara). Randomly selected cloned inserts were sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) with a PRISM 3130 Genetic Analyzer (Applied Biosystems). The 5'-ends of these cDNAs were obtained using the gene-specific primers and the 5'-RACE primer from the CapFishing Full-Length cDNA

Premix Kit. The full-length cDNA clones (CrUGT6 and CrUGT7) were amplified and sequenced using the 5'- and 3'-sequences as specific primers.

3.3.4 Cloning of UGTs by EST database screening

C. roseus EST assemblies on the PlantGDB server (http://plantgdb.org/cgi-bin/blast/PlantGDBblast) were mined using TBLASTN with the GjUGT2 amino acid sequence as a query. From an initial list of UGT candidates obtained by BLASTX search, the contigs belonging to the group G of PSPGs (Nagatoshi *et al*., 2011) were manually selected to provide the master candidate list as shown in supplementary Table 3.S2. We designed gene specific primer and isolated a full-length cDNA corresponding to the contig *Cr8440* as *CrUGT8*.

The nucleotide sequences of PCR primers used for RACE-PCR and amplification of full-length cDNAs of *CrUGTs* were shown in supplementary Table 3.S3 and Table 3.S4, respectively.

3.3.5 Heterologous expression of UGTs

The *CrUGT6*, *7,* and *8* cDNA fragments containing their respective open reading frames were amplified by PCR with KOD FX DNA polymerase (Toyobo, Tokyo) using the PCR primers containing appropriate restriction sites (Table 3.S5). The PCR products were cloned into the pMD20 T-vector to confirm their sequences, and then subcloned into pQE-30 vector (Qiagen) to create N-terminal fusion proteins with a His_6 tag. *Escherichia coli* strain JM109 was used as the host for expression. Transformed cells were cultured at 37 °C until the absorbance at 600 nm reached 0.6 and then further

incubated overnight at 30 °C before harvest. The recombinant proteins were affinity-purified on a Ni-NTA agarose matrix (Qiagen) according to the manufacturer's instruction. Protein content in the enzyme preparations was estimated using a method of Bradford (1976).

3.3.6 Enzyme assays

The standard reaction mixture (total volume of 50 μl) contained 50 mM Tris-HCl (pH 8.0), 5 mM UDP-glucose, 1 mM acceptor substrates, and the enzyme preparation. The reaction was performed at 30 °C and then terminated by the addition of 100 μl of methanol. After centrifugation at 12,000 g for 5 min, the reaction products were analyzed by reversed-phase HPLC, and elution was monitored using a photodiode array detector. To separate the substrates and their glucosides, the following gradient elution programs were used: for iridoid glucosides, 0-16 min, 30–60% methanol and 16–19 min, 60% methanol; for flavonoids, coumarins, and various phenolic glucosides, 0–26 min, 15–52% acetonitrile, 26–29 min, 52–100% acetonitrile, and 29–33 min, 100% acetonitrile. To determine the kinetic parameters, enzyme assays were performed in triplicate at each substrate concentration with 10–25 μg of the purified enzyme at 30 °C for 10 min. The substrate concentrations used were 0.01– 5 mM sugar acceptors with UDP-glucose at 5 mM for acceptor kinetics and 0.25–10 mM UDP-glucose with 7-deoxyloganetin at 0.5 mM for donor kinetics. The initial velocity data were visualized by Lineweaver-Burk plots, and kinetic parameters were calculated based on linear regression analysis using Excel 2007 (Microsoft Japan).

3.3.7 Analysis of gene expression by real-time qRT-PCR

Total RNA was prepared from leaves, stems, flowers and roots of *C. roseus* using the RNeasy Plant Mini Kit. Leaf epidermis-enriched RNA was prepared by the carborundum abrasion technique as described previously (De Luca *et al*., 2012b). First-strand cDNAs for real-time PCR were synthesized from 0.5 µg of total RNA using SuperScript III RNase H^- reverse transcriptase (Invitrogen). Real-time PCR was performed with the 7300 Real Time PCR System (Applied Biosystems), using *Power* SYBR Green PCR Master Mix (Applied Biosystems) according to the manufacturer's instruction. Briefly, the reaction mixture consisted of cDNA template, 10 pmol primers, and 10 µl of *Power* SYBR Green PCR Master Mix in a total volume of 20 µl. The standard PCR condition was as follows: 95°C for 10 min, 40 cycles of 95°C for 15 sec, 60°C for 1 min. Gene-specific primers for CrUGT6, 7, and 8 were listed in Supplementary Table 3.S6.

3.3.8 Virus-induced gene silencing

Virus-induced gene silencing (VIGS) was performed as described previously (Liscombe and O'Connor, 2011) and as modified by De Luca *et al*., (2012b). Fragments of *PDS* (phytoene desaturase) (460 bp) (GenBank accession #: JQ655739), *CrUGT8* (349 bp) (GenBank accession #: KF415118), *LAMT* (loganic acid *O*-methyltransferase) (373 bp) (GenBank accession #: KF415116) and *SLS* (secologanin synthase) (359 bp) (GenBank accession #: KF415117) were amplified by PCR using gene specific primers listed in Table 3.S7. Each amplicon was cloned into the pGEM-T easy vector (Promega) and mobilized to pTRV2 vector after digestion with appropriate restriction enzymes.

Agrobacterium tumefaciens strain GV3101 harboring pTRV1, empty pTRV2 vector (EV) or the pTRV2 construct was cultured overnight at 28°C in 300 mL of Luria-Bertani (LB) medium containing 10 mM MES, 20 μM acetosyringone, and 50 μg mL^{-1} kanamycin. These cultures were centrifuged at 5,000*g* for 10 min and the bacterial pellets were resuspended in 5mL of infiltration buffer (10 mm MES, 200 μM acetosyringone, and 10 mM $MgCl_2$) and further incubated at 28°C for 3 hr with shaking.

C. roseus cv. Little Delicata seeds were germinated and grown in a greenhouse under a 16/8 hr light/dark photoperiod at 28 ºC for 3-6 weeks to produce at least two true leaf pairs. Young plants were wounded using toothpicks through the stem just below the apical meristem and infiltrated with a 1:1 (v/v) mixture of *A. tumefaciens* cultures harboring pTRV1 and either empty pTRV2 vector or pTRV2 constructs. Typically the PDS phenotype was observed 3 week after inoculation of seedlings to signal the stage of growth where leaves from control uninoculated WT, mock-, EV-, *CrUGT8*-, *LAMT*- and *SLS*-inoculations were harvested. After recording fresh weights of harvested materials, one member of a leaf pair was used for RNA extraction, while the other was used for metabolite analysis. Leaf tissues were frozen in liquid nitrogen and were submitted to extraction with a tissue lyser (TissueLyser II, Qiagen) for quick pulverization. Frozen 2 mL microfuge tubes containing leaf materials and 100 μL of 1 mm and 2 mm glass bead mixtures (4:1 ratio) were transferred to a frozen (-80 °C for at least 2 h) TissueLyser Adapter Set that can accommodate 24 samples/plate for performing extractions. Tissue lysis was conducted at 30 Hz for 1 min after which sample were cooled in liquid nitrogen for 1-2 min, and tissue lysis was repeated for another min. Samples were then extracted for RNA or for metabolites.

3.3.9 VIGS-treated plant RNA isolation

Tissue lysed leaf material was mixed with 1 mL Trizol® Reagent (Invitrogen) and were incubated with shaking [100 rpm, Innova 2000 shaker (New Brunswick Scientific)] at room temperature for 5 min. Trizol extracts were mixed with 200 µL of chloroform and after incubated for 3 min at room temperature the phases were separated by centrifugation (15 min at 14,000 rpm at 4°C). The aqueous phase was harvested, mixed with 0.7 volume of isopropanol, incubated for 1 h at room temperature and the RNA pellet was harvested by centrifugation (14,000 rpm at 4°C for 30 min) The pellet was washed with 75% ethanol in DEPC water (1 mL), decanted, dried with SPD Speed Vac Thermo Savant for 2 min, air dried for 5 min and then resuspended in 200 µL DEPC water. The RNA was washed [200 µL of phenol: chloroform: isoamyl alcohol (25:24:1), saturated with 10 mM Tris pH 8.0, 1mM EDTA pH 8.0] by mixing for 3 min and the phases were separated by centrifugation (14,000 rpm at 4°C for 15 min). The aqueous phase was washed with chloroform (200 µL) by mixing for 3 min and the phases were separated by centrifugation (14,000 rpm at 4°C for 15 min). RNA in the aqueous layer was precipitated by overnight incubation at 4°C in the presence of 2.67 M LiCl followed by centrifugation (14,000 rpm at 4°C for 30 min). The pellet was washed with 75% ethanol in DEPC water (1 mL), decanted, dried with SPD Speed Vac Thermo Savant for 2 min, air dried for 5 min and then resuspended in 20 µL MilliQ water. After 10 min at room temperature, 3 µL of each sample was incubated with DNase (1 µL of each 10x DNase buffer (New England BioLabs + 1 µL DNase (New England Biolabs + 5 µL of milliQ water) at 37°C for 30 min. The DNA digestion was stopped by addition of 2 µL of 25 mM EDTA and heating in dry bath at 75°C for 10 min. The samples were cooled

down for 15 min before performing RT-PCR. Single stranded cDNA was synthesized from 2.0 μg RNA using Avian Myeloblastosis Virus (AMV) Reverse Transcriptase (Promega) and oligo (dT) (Alpha DNA) following the manufacturer's protocol. Detection of TRV coat protein in plants infiltrated with pTRV vectors were done as described previously (Rotenberg *et al*., 2006). Primers for detection of TRV coat protein were listed in Table 3.S7.

3.3.10 Metabolite analysis by UPLC-MS

Tissue lysed leaf material was extracted in 1 ml methanol at room temperature with occasional mixing over a 2 hr period. An 200 mL aliquot of the methanol extract was filtered through PALL filter (0.22 μm, VWR) before analysis by UPLC/single-quadrupole MS. Iridoid analysis was carried out using ACQUITY UPLC system (Waters) equipped with an ACQUITY UPLC BEH C_{18} column (1.0 × 50 mm dimension with particle size of 1.7 μm, Waters) and a mobile phase consisting of solvent A (0.1% (v/v) formic acid) and solvent B (100% acetonitrile). Iridoids were eluted at a flow rate 0.3 mL/min with the following linear gradient: 0-0.5 min 99% A, 1% B; 0.5-5.0 min 92% A, 8% B; 5.0-6.5 min 70% A, 30% B; 6.5-7.2 min 50% A, 50% B; 7.2-7.5 min 70% A, 30% B, 7.5-8.0 min 92% A 8% B; 8.0 min 99% A, 1% B. 7-Deoxyloganetic acid, 7-deoxyloganic acid, loganic acid, loganin, and secologanin reference standards were also analyzed by this method and standard curves were generated to measure the levels of each iridoid in the extracts. The mass spectrometer was operated with an ESI ion source of positive ionization mode with a capillary voltage of 3.0 kV, cone voltage of 30 V, cone gas flow of 1 L/hr, desolvation gas flow of 600 L/hr, desolvation temperature of 350°C,

and a source temperature of 150°C. All iridoids were detected at 240 nm with secologanin (m/z = 389, RT = 3.88 min), 7-deoxyloganic acid (m/z = 361, RT = 4.06 min), 7-deoxyloganetic acid (m/z = 197, RT = 4.68 min), loganin (m/z = 391, RT = 3.10 min), and loganic acid (m/z = 377, RT = 1.90 min). Alkaloids were identified and measured as described previously (Roepke *et al*., 2010).

3.3.11 In situ hybridization

The *in situ* RNA hybridization was performed basically as described previously (St-Pierre *et al*., 1999) with some modifications. The first pair of leaves from *C. longifolius* were fixed in FAA (50% ethanol, 5% acetic acid, and 5% formaldehyde), dehydrated through an ethanol and tert-butanol series and then embedded in Paraplast Xtra (Fisher Scientific). The embedded samples were sectioned into 10 µm thickness by using a rotary microtome (Reichert Jung). The sections were carefully spread onto the slides treated with 2% (v/v) 3-aminopropyltriethoxysilane (AES; Sigma) in acetone, incubated for 24 hr at 40°C and stored at 4°C until use. Serial sections were deparaffinized by two incubations of 15 min each in xylene before rehydration in an ethanol gradient series up to diethylpyrocarbonate-treated water.

Full-length of *CrUGT8* in the pGEM-T easy vector (Promega) were used for the synthesis of sense and antisense digoxigenin-labeled RNA probes using DIG RNA Labeling Kit (SP6/T7, Roche) according to the manufacturer's instructions. The RNA probes were submitted to partial alkaline hydrolysis for 20 min at 60 °C. After prehybridization, hybridization of the digoxigenin-labeled RNA probes and washing, the

slides were stained with alkaline phosphatase-conjugated antidigoxigenin antibodies (Roche).

3.4 Results

3.4.1 Molecular cloning of CrUGTs from C. roseus cell cultures and leaves

Total RNA prepared from *C. roseus* cell cultures was used as the template for RT-PCR cloning of CrUGTs using primers based on the conserved amino acid sequence within the PSPG-box. Six partial cDNA fragments were obtained with their deduced amino acid sequences similar to the C-terminal sequences of various PSPGs in the database but not identical with the UGTs previously isolated from *C. roseus* (Kaminaga *et al*., 2004; Masada *et al*., 2009). Using these partial cDNA fragments, we obtained two nearly full-length cDNAs by 5′-RACE and designated them as *C. roseus* UDP-sugar glycosyltransferase 6 (*CrUGT6*) and 7 (*CrUGT7*). In addition, EST-database mining of a *C. roseus* PlantGDB (http://www.plantgdb.org/) database with an iridoid-specific glucosyltransferase from *Gardenia jasminoides* (Nagatoshi *et al*., 2011; *GjUGT2*) identified 30 putative UGT cDNA contigs. Among these, eight contigs (Table 3.S1) associated with group G PSPGs where GjUGT2 belongs were used to isolate an additional full-length cDNA clone (*CrUGT8*) corresponding to contig Cr8440 from leaves of *C. roseus*. An identical contig (CROWL1VD) to Cr8440 was also identified from the PhytoMetaSyn *C. roseus* database (http://www.phytometasyn.ca), but not sequences to *CrUGT6* and *CrUGT7*.

Phylogenetic analyses based on the deduced amino acid sequences of CrUGT6, CrUGT7, and CrUGT8 suggested that while CrUGT6 and CrUGT8 both belonged to

group G, CrUGT7 belonged to group H of Family 1 PSPGs (Figure 3.S1). Functionally characterized members of group G are involved in the biosynthesis of the iridoid geniposide (GjUGT2) in *Gardenia jasminoides*, cyanogenic glucosides (UGT85B1) in *Sorghum bicolor*, cytokinins (UGT85A1) in *Arabidopsis thaliana*, and in an undisclosed reaction in *Lonicera japonica* (LjUGT12). Members of group H are involved in the biosynthesis of steviosides (UGTG1) in *Stevia rebaudiana*, cytokinins (UGTC1) in *Arabidopsis thaliana*, and benzoxazines (*Zea* Bx8) biosynthesis in *Zea mays*. The amino acid sequence comparisons showed 78, 33, and 40 % sequence identity of CrUGT6, CrUGT7, and CrUGT8, respectively with the *Gardenia* iridoid-specific glucosyltransferase (GjUGT2 or UGT85A24). These results might suggest that the best candidate for an iridoid specific GT was CrUGT6.

3.4.2 Functional characterization of recombinant CrUGTs

To examine the catalytic function of CrUGT6-8, their ORFs were expressed in *E. coli* as *N*-terminal fusion proteins with a His_6-tag. After purifying each recombinant CrUGT by Ni-NTA affinity chromatography they were assayed for their *O*-glucosyltransferase activity by using 7-deoxyloganetic acid and 7-deoxyloganetin as acceptor substrates in the presence of the UDP-glucose donor (Figure 3.2). *CrUGT6* rapidly and efficiently converted 7-deoxyloganetin to a product with an identical retention time and UV spectrum as 7-deoxyloganin whereas no reaction product was detected with 7-deoxyloganetic acid as substrate. *CrUGT8* effectively produced 7-deoxyloganic acid from 7-deoxyloganetic acid while no such conversion was detected when 7-deoxyloganetin was used as a sugar accepting substrate. In contrast, small

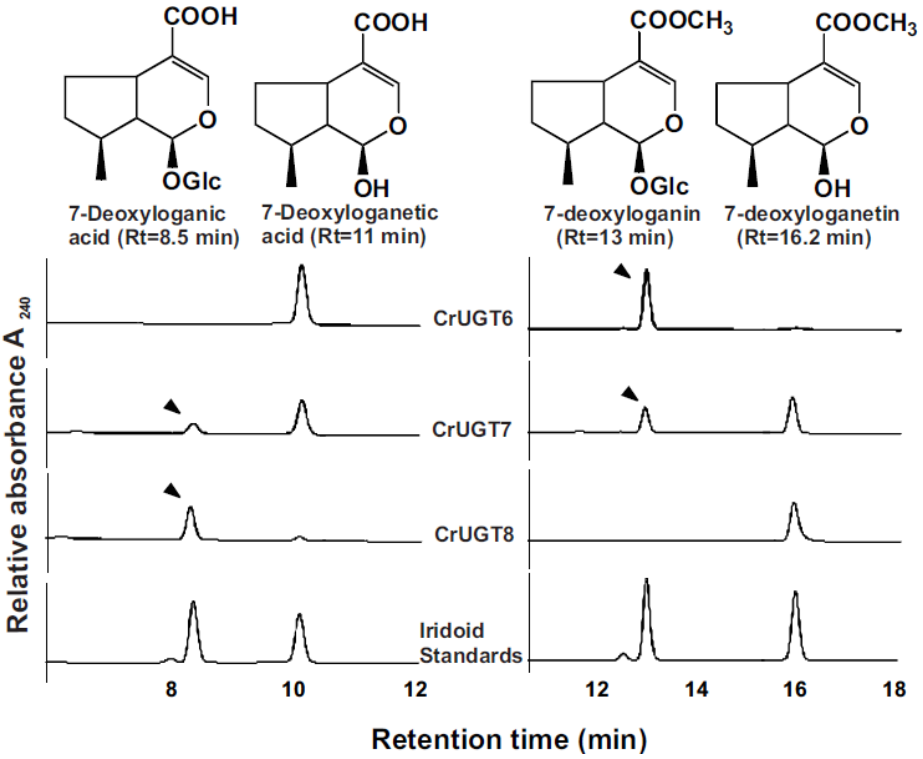


Figure 3.2 Differential conversions of 7-deoxyloganetic acid and 7-deoxyloganetin by recombinant CrUGT6, CrUGT7, and CrUGT8 to 7-deoxyloganic acid and 7-deoxyloganin. Iridoid substrates (1 mM) were incubated with each rCrUGT in the presence of 5 mM UDP-glucose for 2 hrs at 30 °C, and the reaction mixture was subjected to HPLC analysis as described in Methods.

Table 3.1 Kinetic parameters of recombinant CrUGT6-8 toward iridoid aglycones

	Kinetic parameters for 7-deoxyloganetic acid		
Protein	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
CrUGT6	—	—	—
CrUGT7	61.1	0.00164	26.8
CrUGT8	87.7	0.131	1484.1
	Kinetic parameters for 7-deoxyloganetin		
Protein	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
CrUGT6	0.142	0.0115	81.0
CrUGT7	2.12	0.00520	2.45
CrUGT8	—	—	—

Data represent means from triplicate measurements.

amounts of products were detected when *CrUGT7* was incubated with either 7-deoxyloganetic acid or 7-deoxyloganetin.

The glucosyl acceptor specificities of *CrUGT6-8* were tested against a broader range of iridoids, phenolics, flavonoids and one hormone as glucosyl acceptor substrates (Figure 3.S2). *CrUGT6* and *CrUGT8* exhibited strict substrate specificity toward 7-deoxyloganetin and 7-deoxyloganetic acid, respectively. In contrast, *CrUGT7* that is a weak iridoid glucosyltransferase (Figure 3.2) did glucosylate a broader range of substrates including curcumin, genistein, luteolin and kaempferol. Finally, kinetic parameters of *CrUGT6-8* for 7-deoxyloganetic acid and 7-deoxyloganetin were determined based on the pseudo-single substrate kinetics using UDP-glucose as a sugar donor substrate (Table 3.1). The apparent K_m values for 7-deoxloganin of *CrUGT6* and 7-deoxyloganetic acid of *CrUGT8* were 0.142 mM and 0.0611 mM, respectively. The k_{cat}/K_m ratio of *CrUGT8* for 7-deoxyloganetic acid was about 18-fold higher than that of *CrUGT6* for 7-deoxyloganetin. Together these analyses suggest that only *CrUGT8* possessed a high catalytic efficiency toward its exclusive iridoid substrate, 7-deoxyloganetic acid (Table 3.1) and that this enzyme might be better suited for the biosynthesis of iridoids in *C. roseus*.

3.4.3 Preferential expression of CrUGTs in plant organs and leaf cells

The transcript levels of *CrUGT6-8* were determined in *C. roseus* plant organs (Figure 3.3) by real time qRT-PCR and their relative abundance was compared to the levels of secologanin, catharanthine and vindoline (Figure 3.3) present in each organ. Remarkably, expression of *CrUGT6* occurs preferentially in roots, while *CrUGT7*

transcripts were substantially more abundant in leaf pairs 1, 2 and 3 compared to the low levels observed in roots, stems and flowers. The expression of *CrUGT8* transcripts was detected mostly in leaves, roots and stems with lower levels occurring in flowers. The transcript level of *CrUGT8* was highest in the youngest leaves (Figure 3.3, leaf pairs 1 to 3) and was gradually decreased in the older leaves (Figure 3.3, leaf pairs 4 and 5). The patterns of expression of *CrUGT8* were very similar to those of *LAMT* and *SLS* (Figure 3.3) that is consistent with the significant levels of secologanin found in most *C. roseus* organs (Figure 3.3). The results also confirm that only part of the secologanin produced is incorporated into MIAs that are preferentially biosynthesized in younger tissues (leaf pairs 1 and 2 and in root tips; Murata *et al*., 2008; Roepke *et al*., 2010) The accumulation of MIAs (catharanthine and vindoline) reaches a maximum by leaf pair 3 (Figure 3.3) while secologanin continues to accumulate until leaf pair 4 (Figure 3.3, see mg/organ). Similarly the levels of secologanin increase at least 10-fold in fully open flowers compared to flower buds (Figure 3.3, see mg/organ).

Carborundum abrasion technique has been successfully used to extract leaf epidermis enriched RNA from *C. roseus* and to corroborate MIA pathway gene expression in these cells (Levac *et al*., 2008; Murata *et al*., 2008). qRT-PCR analysis of leaf epidermis enriched transcripts showed that the levels of *CrUGT6* and *CrUGT8* were at least 10-fold lower than those found in whole leaves, while *CrUGT7* transcripts were more equally distributed (Figure 3.4A). In contrast transcripts for loganic acid methyltransferase (*LAMT*) and secologanin synthase (*SLS*), well known to be preferentially expressed in the leaf epidermis (Guirimand *et al*., 2011a; Murata *et al*., 2008), were 2- to 4-fold more enriched in these cells. Together these results suggested

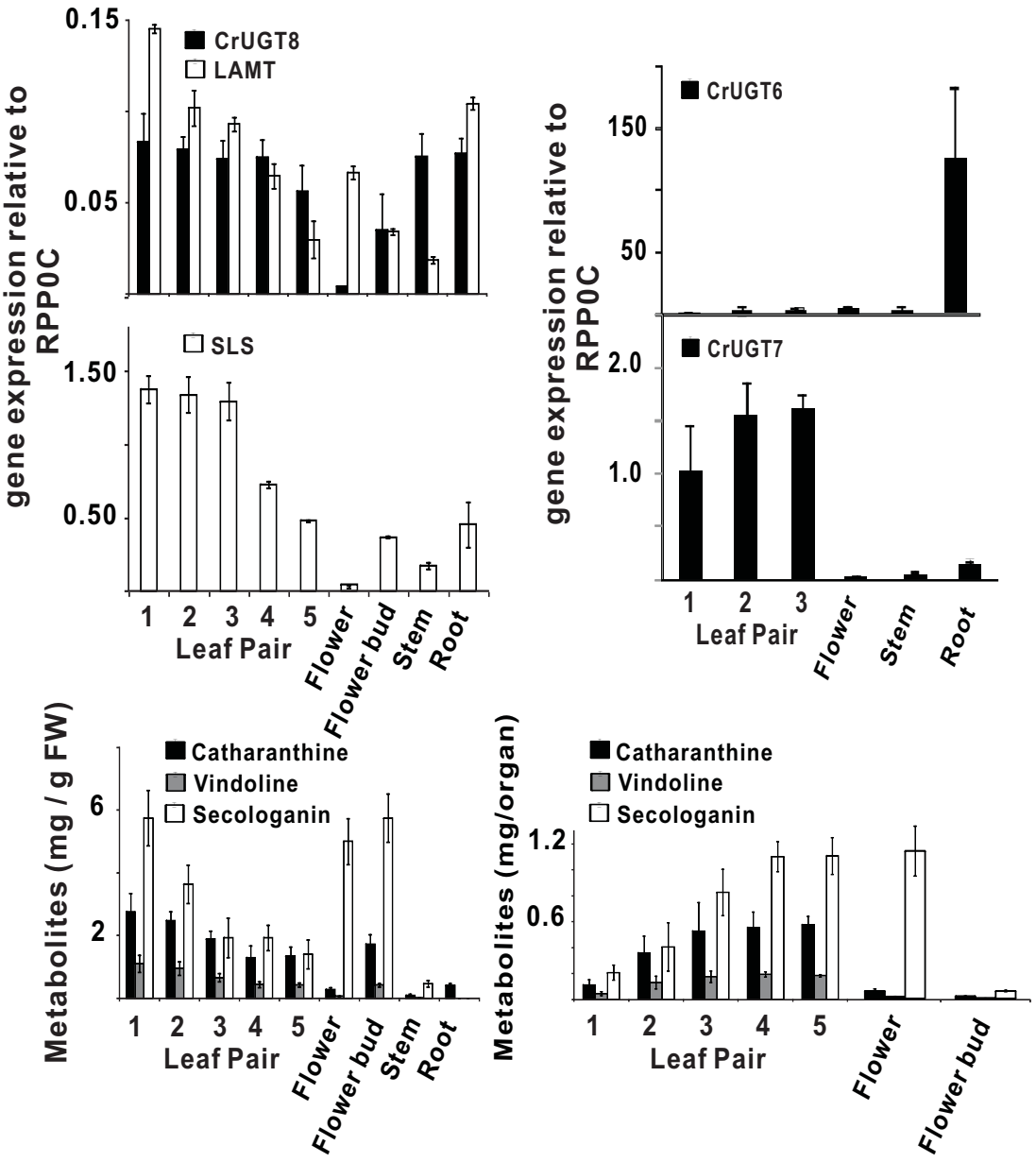


Figure 3.3 Differential expressions of CrUGT8 correlate with those of last 2 steps in secologanin biosynthesis together with iridoid and MIA metabolite profiles in *C. roseus* plant organs. Relative gene expression of *CrUGT6*, *CrUGT7*, *CrUGT8*, *LAMT* and *SLS* were determined by real-time quantitative RT-PCR analyses performed on total RNA extracted from *C. roseus* leaf pairs 1 to 5, from open flowers and flower buds and from root tissues. Each point represents the mean of relative transcript abundance to RPP0C (gene encoding 60s acidic ribosomal protein P0-C) ± standard deviation from triplicate measurements of biological and technical replicates. Metabolite (catharanthine, vindoline and secologanin) levels are plotted as mg/gram fresh weight and as mg/organ. Each box and bar represents an average value and a standard error, respectively, from 4 different plant samples.

that both *CrUGT-6* and *-8* are preferentially expressed within *Catharanthus* leaves rather than in leaf epidermal cells where the last 2 steps in secologanin assembly are expressed.

3.4.4 CrUGT8 is preferentially expressed in internal phloem-associated parenchyma of Catharanthus leaves

The high catalytic efficiency of CrUGT8 toward its exclusive iridoid substrate, 7-deoxyloganetic acid (Table 3.1) and its expression within *Catharanthus* leaves rather than in leaf epidermis prompted *in situ* hybridization studies to localize where this gene is preferentially expressed. Very young developing leaves of *C. roseus* and *C. longifolius* were harvested, fixed, embedded, sectioned, and prepared for *in situ* RNA hybridization analysis to localize transcripts of *CrUGT8* (Figure 3.4). Remarkably, labeling with antisense *CrUGT8* probes was restricted to the adaxial phloem region in *C. longifolius* tissues surrounding the vasculature on longitudinal leaf sections (Figure 3.4B and 3.4D), while no significant labeling was observed when sense *CrUGT8* probes were used in negative control experiments (Figure 3.4C and 3.4E). Unfortunately, the same experiments performed with *C. roseus* did not produce any labeling of the same sections, perhaps due to the several-fold lower abundance of *CrUGT8* transcript found compared to the levels occurring in young leaves of *C. longifolius*. It should be noted that the sequence of CloUGT8 is 100% identical to that of CrUGT8 (Figure 3.S1). These results strongly suggested that CrUGT8 is expressed in internal phloem-associated parenchyma cells that also preferentially express the 2-C-methyl-D-erythritol 4-phosphate pathway (Burlat *et al*., 2004), geraniol synthase (Simkin *et al*., 2013), geraniol-10 hydroxylase (G10H) (Burlat *et al*., 2004), and iridoid synthase (Geu-Flores *et al*., 2012).

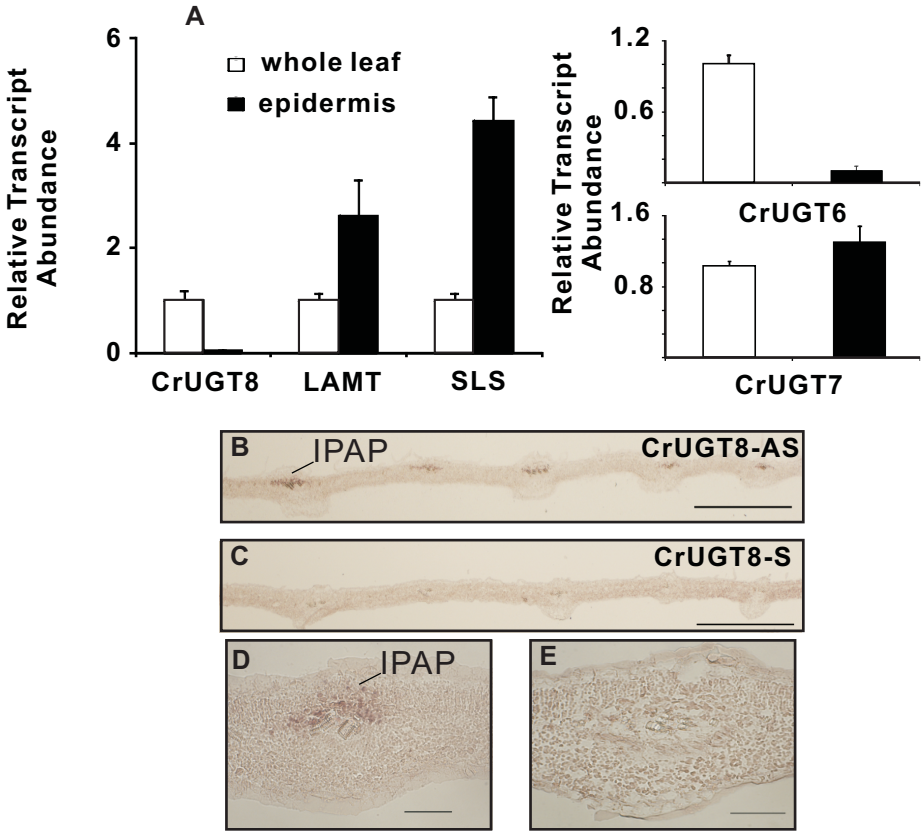


Figure 3.4 Localization of CrUGT8 transcripts in IPAP cells of young developing leaves of *Catharanthus*. A) Relative expression of *CrUGT6*, *CrUGT7*, *CrUGT8*, *LAMT,* and *SLS* in relation to the *CrRPP0C* (encoding 60s acidic ribosomal protein P0-C) reference gene in leaf epidermis enriched transcript extracted by carborundum abrasion compared with those found in whole leaf extracts. **B-E**) Localization by *in situ* hybridization of *CrUGT8* mRNA in young developing leaves of *Catharanthus longifolius*. Serial longitudinal 10 µm sections were made from young leaves (10 to 15 mm long) were hybridized with *CrUGT8*-antisense (AS) (**B, D**); *CrUGT8*-sense (S) (**C, E**) probes. Bars in **B** and **C** are 500 µm while in **D** and **E** they are 50 µm.

3.4.5 Virus induced gene silencing of CrUGT8, LAMT, and SLS trigger large declines in accumulation of secologanin and MIAs in silenced Catharanthus leaves

The strict specificity of CrUGT8 for 7-deoxyloganetic acid (Figure 3.2 and Figure 3.S2), its higher catalytic efficiency than the other CrUGTs (Table 3.1), its coordinate expression with *LAMT* and with *SLS* together with the accumulation of metabolites in relevant *Catharanthus* tissues (Figure 3.3), and its preferred expression in IPAP cells of *Catharanthus* leaves (Figure 3.4) prompted us to use virus-induced gene silencing (VIGS) for validating the role of CrUGT8 in iridoid biosynthesis in *C. roseus* (Figure 3.5). Fragments of *CrUGT8*, *LAMT* and *SLS* genes (Table 3.S7) were cloned independently into pTRV2 vectors. A mixed culture of *Agrobacterium tumefaciens* harboring pTRV1 and either pTRV2 vector (pTRV2-EV) or pTRV2 constructs (pTRV2-CrUGT8, pTRV2-LAMT, pTRV2-SLS) was infiltrated into the apical meristem of *C. roseus* young plants (De Luca *et al*., 2012b). Separate experiments suppressing the phytoene desaturase (PDS) gene (Table 3.S7) produced a visible photobleached phenotype that was monitored as a visible marker for determining the timing to perform transcript and metabolite analyses in plants suppressed for *CrUGT8*, *LAMT* and *SLS*. In order to confirm the success of *Agrobacterium* infiltration, plants were selected based on the detection of a 134 bp fragment derived from the presence of the pTRV2-derived TRV coat protein transcript (Figure 3.S3). *Catharanthus* plants suppressed for each transcript showed large declines in the levels secologanin accumulation as determined by UPLC-MS (Figure 3.5A). Plants suppressed for *CrUGT8* did not show an increase in detectable 7-deoxyloganetic acid (Figure 3.5A) while those suppressed for *LAMT* or for *SLS* both

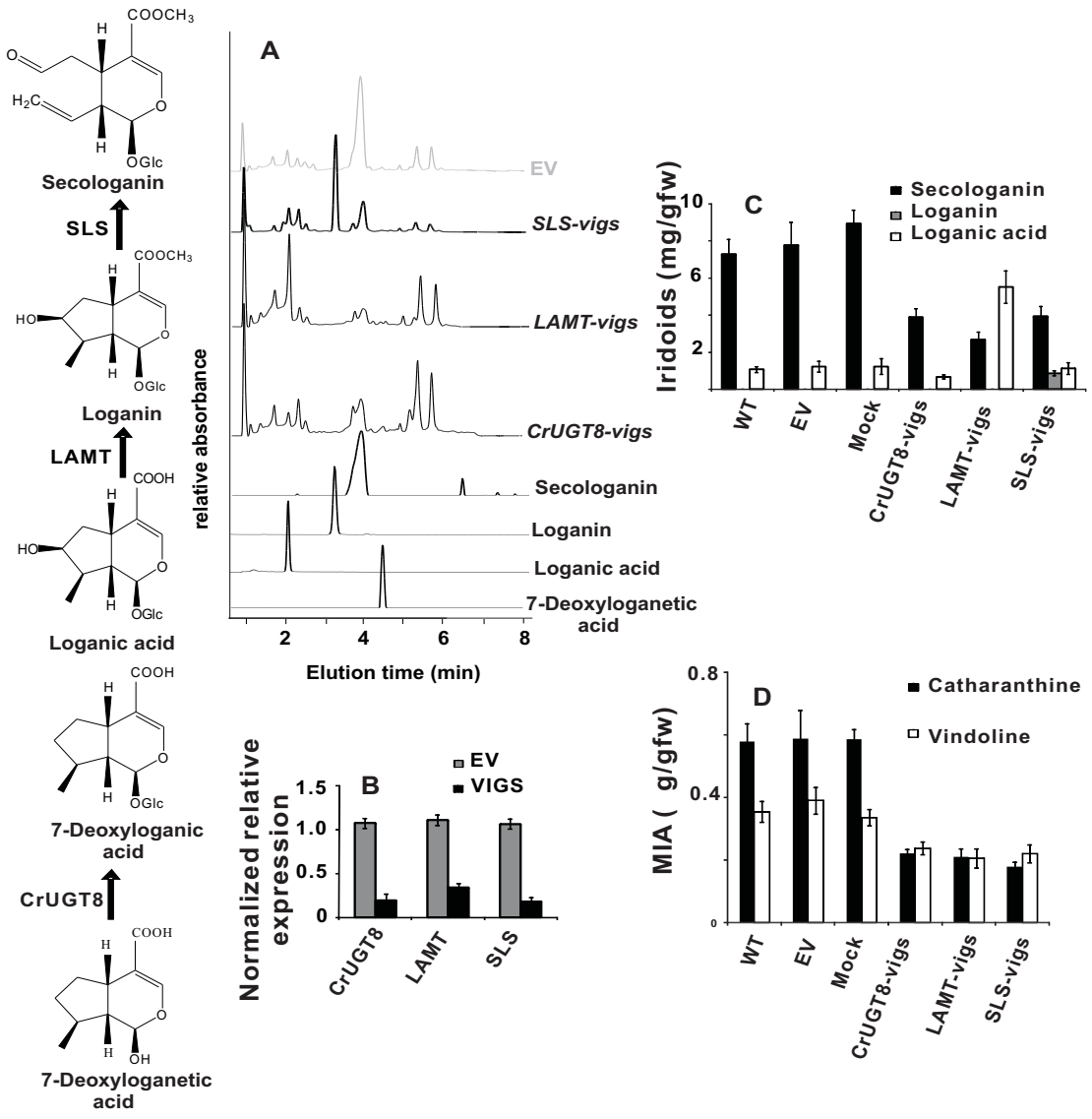


Figure 3.5 Down-regulation of *CrUGT8*, *LAMT,* and *SLS* affect the accumulation of iridoids and MIAs in *C. roseus*. A) Silencing of *CrUGT8*, *LAMT* and *SLS* was conducted by monitoring the iridoid metabolite profiles by UPLC-MS in silenced plants (*UGT8-vigs*, *LAMT-vigs*, and *SLS-vigs*) compared to the profiles obtained with plants treated with empty vector (EV) controls. UPLC-MS analysis of iridoid profiles were detected at A_{240} nm and by retention times related to iridoid standards [7-deoxyloganetic acid (RT = 4.68 min ; m/z = 197), loganic acid (RT = 1.90 min; m/z = 377), loganin (RT = 3.1 min; m/z =391), secologanin (RT = 3.88 min; m/z = 389). **(B)** Silencing of *CrUGT8*, *LAMT*, and *SLS* was measured by monitoring relative transcript abundance of each iridoid pathway gene by real-time quantitative RT-PCR. Differences in transcript levels for each silenced gene were measured relative to those obtained in EV and Mock treatments and are represented as mean ± standard error. Gene specific primers for each *CrUGT8*, *LAMT*, and *SLS* were used for comparison of transcript abundance between EV and for each VIGS treatment. The data represents measurements performed with six biological replicates (with 3 technical replicates per biological replicate) of Mock, EV, *CrUGT8-vigs*, *LAMT-vigs*, and *SLS-vigs* treatments. Measurements of iridoids (loganic acid, loganin, secologanin) **(C)** and MIAs (catharanthine and vindoline) **(D)** in untreated (WT), EV, Mock, *CrUGT8-vigs*, *LAMT-vigs*, and *SLS-vigs* treated *Catharanthus* plants were performed with the same 6 biological replicates used for transcript analysis in **B**. Significant differences were considered with P values of <0.05 by Student's *t*-test for the transcript analysis and metabolite contents of EV-infected plants and in each of the silenced lines.

showed large increases in detectable loganic acid and loganin accumulation, respectively. These results were verified by monitoring relative transcript levels of *CrUGT8*, *LAMT* and *SLS* by quantitative RT-PCR. The transcript levels declined by 70 to 80 % (P value of <0.001) in *CrUGT8*-, *LAMT*- and *SLS*-vigs leaf tissues compared with EV controls (Figure 3.5B). More detailed metabolite analyses of *CrUGT8*-, *LAMT*- and *SLS*-silenced tissues showed that they responded with more than 50 % decline of secologanin accumulation (Figure 3.5C) together with variable increases in loganic acid levels and only *LAMT*-silenced tissues responded with accumulation of loganic acid, which is the substrate for *LAMT* (Figure 3.5C) compared to *Catharanthus* control, empty vector and mock inoculated tissues. More detailed MIA analyses of *CrUGT8*-, *LAMT*- and *SLS*-silenced tissues revealed that they all responded with more than 50 % decline of catharanthine accumulation (Figure 3.5D) together with smaller 30 to 40 % declines in vindoline levels compared to *Catharanthus* control, empty vector and mock inoculated tissues.

3.5 Discussion

3.5.1 C. roseus contains CrUGTs that glucosylate iridoids with varying efficiencies

Three functional CrUGT clones were obtained from cell suspension culture (*CrUGT6* and *CrUGT7*) and leaf (*CrUGT8*) RNA preparations of *Catharanthus* based on the conserved PSPG-box of this family of glucosyltransferases and sequence information from the only characterized iridoid glucosyltransferase from *Gardenia jasminoides* (Nagatoshi *et al*., 2011; *GjUGT2*). While *CrUGT6* and *CrUGT8* belong to Group G of Family 1 PSPGs (Figure 3.S1), *CrUGT7* aligns with group H of this family. The

biochemical function of CrUGT6 whose amino acid sequence is 78 % identical to GjUGT2 is most similar to the *Gardenia* enzyme that also shows strict specificity for 7-deoxyloganetin rather than 7-deoxyloganetic acid. In contrast, CrUGT8 whose amino acid sequence is 40 % identical to GjUGT2 only used 7-deoxyloganetic acid as a substrate and this enzyme possessed a high catalytic efficiency for its substrate compared to the other identified iridoid GTs (Table 3.1). Inspection of the PhytoMetaSyn *C. roseus* database (http://www.phytometasyn.ca) also identified genes (Figure 3.S1) that were identical, *Catharanthus ovalis* (CovUGT8), *Catharanthus longifolius* (CloUGT8) or corresponding closely to CrUGT8 from *Tabernaemontana elegans* (TeUGT8), *Amsonia hubrichtii* (AhUGT8), *Rauwolfia serpentina* (RsUGT8), and *Vinca minor* (VmUGT8). In addition, similar candidate genes were readily identified in databases of *Lonicera japonica* (LjUGT8) that makes secologanin and *Cinchona ledgeriana* (ClUGT8) that makes quinoline alkaloids. The presence of such highly similar UGTs in each species of plants that produce iridoids and MIAs provides additional support for the key role played by CrUGT8 in secologanin biosynthesis in *C. roseus* and more generally members of the Apocynaceae family.

3.5.2 Transcriptional down regulation of CrUGT8 by VIGS suppresses secologanin and MIA accumulation in Catharanthus plants

VIGS technology has been successfully exploited together with modern bioinformatic approaches to select and identify candidate genes that may be involved in the biosynthesis of particular metabolites in plants (De Luca *et al*., 2012a). The use of VIGS has been validated in *C. roseus* where VIGS-mediated suppression of 16-methoxy-

2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase (NMT) resulted in plants that stopped accumulating vindoline in favor of 16-methoxy-2,3-dihydro-3-hydroxytabersonine (Liscombe and O'Connor, 2011). Remarkably, VIGS suppression of iridoid cyclase also led to plants containing reduced levels of catharanthine and vindoline (Geu-Flores *et al*., 2012). In the present study, the silencing of *CrUGT8*, *LAMT*, and *SLS* decreased the levels of each respective transcript by 70 to 80 % (Figure 3.5B) and to general declines (Figure 3.5D) of >50 % in cantharanthine and 30 to 40 % of vindoline levels in silenced leaves. Silencing of *CrUGT8*, *LAMT*, and *SLS* also greatly decreased the levels of secologanin by more than 50 % in silenced lines (Figure 3.5C). Remarkably, silencing of *LAMT* and *SLS* resulted in the accumulation of considerable amounts of their respective substrates loganic acid (5.21 mg/gFw) and loganin (0.77 mg/gFw) (Figure 3.5C). This clearly illustrates how such silencing events could be used to supply substrates that may not be available commercially or that may be difficult to produce through organic chemistry.

3.5.3 CrUGT8 catalyzes the 4th to last step in secologanin biosynthesis within IPAP cells of Catharanthus

Many of the steps involved in secologanin biosynthesis from the MEP pathway have been identified at the biochemical and molecular level including geraniol synthase (GES), G10H (Figure 3.1A), a putative 10-hydroxygeraniol oxidoreductase (Figure 3.1B), iridoid synthase (Figure 3.1C), LAMT (Figure 3.1G) and SLS (Figure 3.1H). The most recent molecular and biochemical characterization of iridoid synthase showed it to be a unique progesterone 5β-reductase-like gene that catalyzed a unique reductive cyclization

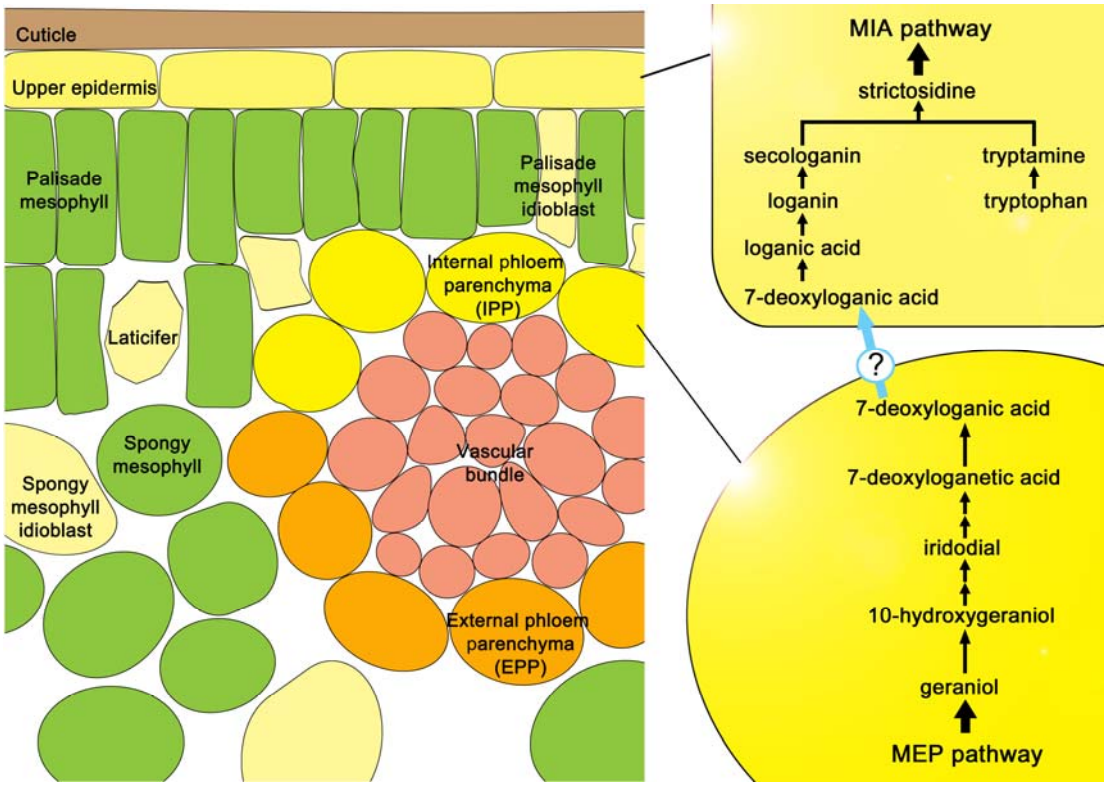


Figure 3.6 Spatial model of iridoid biosynthesis and translocation in *Catharanthus* leaves. The MEP pathway and iridoid biosynthesis to 7-deoxyloganic acid occurs in IPAP cells while the terminal LAMT and SLS reactions occur in the leaf epidermis. The model shows 7-deoxyloganic acid hydroxylase in leaf epidermis, but this reaction and its location still remains to be elucidated. The model shows that MIA assembly from secologanin and tryptamine also takes place in leaf epidermal cells. Solid lines represent a single enzymatic step, whereas double arrows indicate the involvement of multiple enzyme steps. The '?' indicates the putative transport system of 7-deoxyloganic acid from IPAP cells to leaf epidermal cells.

in iridoid biosynthesis and that was preferentially expressed in IPAP cells (Geu-Flores *et al*., 2012).

The present study provides strong evidence that CrUGT8 catalyzes the 4th to last step in secologanin biosynthesis (Figure 3.1, Reactions A to H) and is also preferentially expressed within IPAP cells (Figure 3.4). Since nothing is known about the hydroxylation step that converts 7-deoxyloganic acid to loganic acid, the model shown in Figure 3.6 suggests that this metabolite may be transported to the leaf epidermis to be converted into

loganic acid followed by methylation (LAMT) and oxidative ring opening (SLS) in the leaf epidermis. The molecular cloning of the hydroxylase responsible for the 3rd to last step in secologanin biosynthesis should provide tools to identify its presence in IPAP or in epidermal cells of *Catharanthus* leaves.

3.6 Conclusions

In conclusion, three separate CrUGTs with distinct substrate specificities and catalytic efficiencies have been described in the present study. The optimal catalytic and biochemical properties of CrUGT8, its preferred expression in IPAP leaf cells of *Catharanthus* together with the loss of iridoids and MIAs in VIGS silenced plants strongly suggest its key role as a biosynthetic enzyme in the assembly of secologanin. These combined approaches of using bioinformatics, silencing technologies and functional characterization of candidate genes is revolutionizing the gene discovery process for mining the chemical diversity of plants (De Luca *et al*., 2012a; Facchini *et al*., 2012; Góngora-Castillo *et al*., 2012; Xiao *et al*., 2013).

3.7 Acknowledgements

The authors thank Prof. Peter J. Facchini, University of Calgary, for pTRV1 and pTRV2, Prof. Kenichiro Inoue, Yokohama College of Pharmacy, for 7-deoxyloganin and Dr. Nobuki Kato, Nagoya City University, for iridotrial. We recognize the skilled technical work of next-generation sequencing personnel at the McGill University-Genome Québec-Innovation Centre. We are grateful to Dr. Christoph Sensen, Mei Xiao, and Ye Zhang for their dedicated bioinformatic support and large scale gene annotation

efforts that helped in the identification of various UGT8 genes from the PhytoMetaSyn website. This work was supported by JSPS KAKENHI Grant Numbers 22710218 (to K.T.), 22590009 and 25460127 (to H.M.). This work was also supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (V.D.L.), Genome Canada, Genome Alberta, Genome Prairie, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners. S.M.A is the recipient of postdoctoral funding from Genome Canada and the Ontario Ministry of Research and Innovation.

Chapter 4

Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid 7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis

This manuscript has been accepted for publication in *The Plant Journal*

Virus induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid 7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis

Vonny Salim[a], Fang Yu[a], Joaquín Altarejos Caballero[b], and Vincenzo De Luca[a]

[a]Department of Biological Sciences, Brock University, St. Catharines, ON, Canada L2S 3A1; [b]Departmento de Química Inorgánica y orgánica, Facultad de Ciencias Experimentales, Universidad de Jáen, 23071, Jáen, Spain

4.1 Abstract

Iridoids make up a major group of biologically active molecules present in thousands of plant species and one versatile iridoid, secologanin is a precursor for the assembly of thousands of monoterpenoid indole alkaloids (MIAs) as well as a number of quinoline alkaloids. This study uses bioinformatics to screen large databases of annotated transcripts from various MIA producing plant species to select candidate genes that may be involved in iridoid biosynthesis. Virus-induced gene silencing of the selected genes combined with metabolite analyses of silenced plants is then used to identify the 7-deoxyloganic acid 7-hydroxylase (CrDL7H) that is involved in the 3rd to last step in secologanin biosynthesis. The silencing of CrDL7H reduced secologanin levels by at least 70 % and increased the levels of 7-deoxyloganic acid to over 4 mg/leaf gram fresh weight compared to control plants where this iridoid is not detected. Functional expression of this CrDL7H in yeast confirmed its biochemical activity and substrate specificity studies showed its preference for 7-deoxyloganic acid over other closely

related substrates. Together these results suggest that hydroxylation precedes carboxy-*O*-methylation in the secologanin pathway in *Catharanthus roseus*.

4.2 Introduction

Catharanthus roseus is the sole source of the important anticancer drugs, vinblastine and vincristine, monoterpenoid indole alkaloids (MIAs) that have been extensively used for treatment of leukemia and lymphoma (van der Heijden *et al*., 2004). This medicinal plant accumulates a number of MIAs that originates from the coupling of the indole typtamine, formed by tryptophan decarboxylase (TDC) and the iridoid secologanin by strictosidine synthase, to form strictosidine from which most other MIAs are generated (O'Connor and Maresh, 2006; El-Sayed and Verpoorte, 2007; Oudin *et al*., 2007a).

Iridoids are cyclopentan-[C]-pyran monoterpenoids that occur widely in numerous medicinal plants, usually as biologically active glucosides with antihepatotoxic, antitumor, antiviral, anti-inflammatory, and other properties (Dinda, Debnath, and Harigaya, 2007; Dinda, Debnath, and Banik, 2011). In plants, iridoids also function as important defense compounds in plant-herbivore interactions through their anti-feeding and anti-microbial activities (Peñuelas et al., 2006). Because of the pharmacological and economic importance of iridoids, there has been intense interest to understand their biosynthetic mechanisms in order to develop biotechnological approaches to their targeted production. While several genes involved in the biosynthesis of secologanin have been cloned and functionally characterized in medicinal *C. roseus*, several genes

remain to be characterized in order to complete this pathway (O'Connor and Maresh, 2006; El-Sayed and Verpoorte, 2007; Oudin *et al*., 2007a).

The first committed step in secologanin biosynthesis (Figure 4.1) involves the oxidation of geraniol at its C-10 position by a cytochrome P450 monooxygenase, geraniol 10-hydroxylase [CYP76B6 (G10H)] (Collu *et a*l., 2001). Further oxidation of 10-hydroxygeraniol to the dialdehyde, 10-oxogeranial, is catalyzed by 10-hydroxygeraniol oxidoreductase (10-HGO) (Ikeda *et al*., 1991). Conversion of this dialdehyde to iridodial occurs by a unique reductive cyclization carried out by an NADPH-dependent 10-oxogeranial cyclase named iridoid synthase (IRS) (Geu-Flores *et al*., 2012). The conversion of iridodial to secologanin involves several enzymes that include possible cytochrome P450-dependent reactions and the precise biochemical reactions have yet to be characterized.

The late stage of secologanin biosynthesis from iridodial involves a three step oxidation to form 7-deoxyloganetic acid, most probably catalyzed by cytochrome P450, followed by glucosylation to produce 7-deoxyloganic acid before leading to two putative ways for loganin synthesis that is further converted to secologanin by secologanin synthase (SLS) by a cytochrome P450 dependent enzyme, CYP72A1 (Uesato et al., 1986; Yamamoto *et al*., 2000; Irmler *et al*., 2000). In the first pathway, a hydroxylation step precedes a methylation step where 7-deoxyloganic acid is hydroxylated by 7-deoxyloganic acid 7-hydroxylase (DL7H) to produce loganic acid, which is then methylated into loganin by loganic acid *O*-methyltransferase (LAMT) that has been cloned and characterized in *C. roseus* (Madyastha *et al*., 1973; Murata *et al*., 2008). This methyltransferase is found to accept loganic acid and secologanic acid, but not

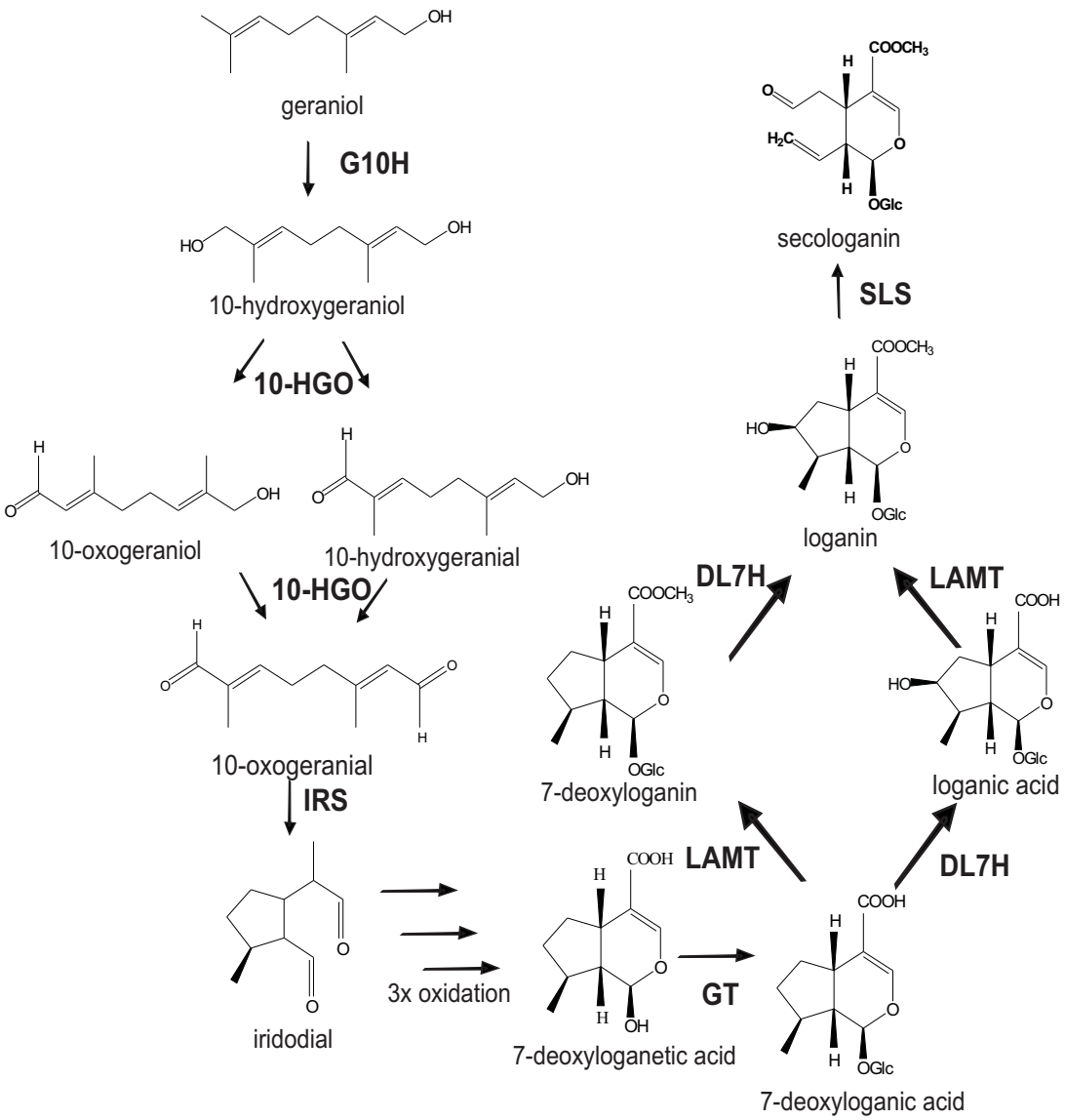


Figure 4.1 Iridoid biosynthesis pathway from geraniol to secologanin. Abbreviations: G10H-geraniol 10-hydroxylase, 10HGO-10-hydroxygeraniol oxidoreductase, IRS-iridoid synthase, GT-glucosyltransferase, DL7H-7-deoxyloganic acid 7-hydroxylase, LAMT-loganic acid *O*- methyltransferase, SLS-secologanin synthase. Multiple arrows indicate multiple steps

7-deoxyloganic acid as a substrate. In the second way, the methylation step precedes hydroxylation in which 7-deoxyloganic acid is methylated to 7-deoxyloganin, which is hydroxylated into loganin. A 7-deoxyloganin 7-hydroxylase has been detected in microsomal preparation of cell suspensions of *C. roseus* as well as *Lonicera japonica* cell cultures with requirement of NADPH and oxygen suggesting the involvement of cytochrome P450 (CYP) for this reaction (Irmler *et al*., 2000; Katano *et al*., 2001). Although this enzyme has been partially purified from cell suspension cultures, no genes encoding the relevant hydroxylase have been cloned and characterized.

The identification of CrDL7H was initiated by searching our EST database of *C. roseus* for a number of cDNAs that encodes putative CYPs with high similarities to those found in other MIA producing plant species and that were also found in the secologanin producing species, *Lonicera japonica* (http://www.phytometasyn.ca/) (Table 4.S1). The PhytoMetaSyn project contains large annotated transcript databases from 75 non-model plants known to produce biologically active terpenes, alkaloids, and polyketides (Facchini *et al*., 2012; Xiao *et al*., 2013). This search for putative CYP genes has led to the identification of three cytochrome P450 candidates that may be involved in iridoid biosynthesis (Table 4.S1). Virus-induced gene silencing (VIGS) of one candidate CYP belonging to the CYP72A subfamily resulted in silenced plants that accumulated 7-deoxyloganic acid rather than secologanin. Recombinant expression and characterization of this CYP in yeast showed it to be a 7-deoxyloganic acid 7-hydroxylase whose order precedes the *O*-methylation step catalyzed by LAMT in secologanin biosynthesis of *C. roseus*.

4.3 Methods

4.3.1 Plant materials

Catharanthus roseus (L.) Don cv. Little Delicata seeds were germinated in 50 well trays, on wet soil in the greenhouse using a 16/8-h light/dark photoperiod at 28°C. For VIGS treatment, the seedlings were grown for 4 weeks or until they produce at least two true leaf pairs.

4.3.2 Iridoid and MIA extraction

Plant tissues harvested for iridoid extraction were frozen and pulverized in liquid nitrogen. Methanol (5 mL) was added and the leaves were homogenized. The samples were placed on an Innova 2000 shaker (New Brunswick Scientific) for 1 hour before metabolite analysis by UPLC-MS. The MIA extraction and analysis were performed according to methods as previously described (Roepke *et al*., 2010).

4.3.3 Chemicals

The amount of 7-deoxyloganic acid (14 mg) used in this study was purified from an ethyl acetate extract of olive wood as follows: a sample of olive wood shavings (200 g) was extracted with CH_2Cl_2 under argon in the dark for 2 h at reflux. The CH_2Cl_2 extract was filtered and concentrated to give a dried residue (1.20 g). Then, the remaining plant material was extracted with ethyl acetate under argon in the dark for 2 h at reflux. The ethyl acetate extract was evaporated under vacuum at temperature not higher than 40°C. The resulting dried ethyl acetate extract (3.20 g) was column chromatographed (silica gel 60, 200 g, 63-200 μm) using CH_2Cl_2-EtOH mixtures of increasing polarity. Fractions of 75 mL each were collected, monitored by thin layer chromatography (TLC)

and high-performance liquid chromatography (HPLC), pooled and evaporated to give ten major fractions (from A-1 to A-10). Fraction A-8 (290 mg) was column chromatographed again (silica gel 60, 40g, 40-63 μm) with the same mixture of solvents as above. Fractions of 25 mL each were collected, monitored by TLC and HPLC, pooled and evaporated to give ten major fractions (from B-1 to B-10). Fraction B-5 (42 mg), enriched in 7-deoxyloganic acid, was re-purified by preparative HPLC to yield pure 7-deoxyloganic acic (14 mg), after evaporating MeOH with a rotary evaporator and the remaining H_2O with a freezer-drier. This final HPLC purification was performed with a preparative reversed-phase HPLC column (Spherisorb ODS-2, 250 mm x 10 mm i.d., 5μm) on a Waters 600E instrument equipped with a diode array detector (scan range, 190-800 nm), operating at 30°C. The separation was carried out by a step gradient with MeOH/AcOH, 99.8:0.2, v/v (solvent A) and H_2O/AcOH, 99.8:0.2, v/v (solvent B) at a flow rate of 5 mL/min: isocratic conditions of 10% A for 5 min; linear gradient from 10% to 100% A in 35 min; isocratic conditions of 100% for 5 min, and other 5 min to return to the initial conditions. The ^1H NMR spectrum of 7-deoxyloganic acid, recorded at 400 MHz on a Bruker Avance 400 spectrometer (using CD_3OD as solvent and tetramethylsilane as internal reference), agrees with that reported for this compound in the literature (Pérez-Bonilla *et al*., 2011). The solvents used for extraction and chromatographic separations were glass-distilled prior to use or purchased HPLC grade. The 7-deoxyloganin and 7-deoxyloganetic acid were kindly provided by Hajime Mizukami, Nagoya City University, Japan as previously described (Chapter 3). Loganic acid, loganin, and secologanin were obtained from Extrasynthese (Genay, France).

4.3.4 Cloning and construction of VIGS vectors

A Fragment of 363 bp from CrDL7H gene was amplified with Ex-Taq polymerase (Takara) and inserted to a pGEMT-easy vector (Promega) using the following primers: forward (5'-CCAAGGCTCTTTTTTTCCAT-3') and reverse (5'-AGAATCCAGTATCCCCAAAT -3'). The sequence was inserted into the multiple cloning sites (MCS) between CaMW 35S promoter (2x35S) and NOS terminator of pTRV2 using a single restriction site, EcoRI, digested from pGEMT-easy vector and ligated into the pTRV2 vector. The orientation of the insert was verified by sequencing. The vector construction for PDS silencing was previously described (De Luca *et al*., 2012b) using the following: a forward primer, 5'-AGGTTTGGGGGGTTTGTGT-3' and a reverse primer, 5-' TACGCCTTGCTTTCTCATCC-3'. Assembly of the pTRV1 and pTRV2 vectors were performed as previously described (Dinesh-Kumar *et al*., 2003).

4.3.5 Virus-induced gene silencing

Virus-induced gene silencing (VIGS) was performed as described previously (Liscombe and O'Connor, 2011) and as modified by De Luca *et al*., 2012b. *Agrobacterium tumefaciens* strain GV3101 harboring pTRV1, empty pTRV2 vector (EV) or the pTRV2 construct was cultured overnight at 28°C in 50 mL of Luria-Bertani (LB) medium containing 10 mM MES, 20 μM acetosyringone, and 50 μg mL^{-1} kanamycin. These cultures were centrifuged at 5,000*g* for 10 min and the bacterial pellets were resuspended in 5 mL of infiltration buffer (10 mm MES, 200 μM acetosyringone, and 10 mM $MgCl_2$) and further incubated at 28°C for 3 hr with shaking.

C. roseus cv. Little Delicata seeds were germinated and grown in a greenhouse under a 16/8 hr light/dark photoperiod at 28°C for 3-6 weeks to produce at least two true leaf pairs. Young plants were wounded using toothpicks through the stem just below the apical meristem and infiltrated with a 1:1 (v/v) mixture of *A. tumefaciens* cultures harboring pTRV1 and either empty pTRV2 vector or pTRV2 constructs. Typically the PDS phenotype was observed 3 weeks after inoculation of seedlings and at this stage, leaves from control uninoculated plants, mock-, EV-, CrDL7H-inoculations were harvested. After recording fresh weights of harvested materials, one member of a leaf pair was used for RNA extraction, while the other was used for metabolite analysis. The extraction of RNA and metabolites from VIGS-treated plants was previously described in Chapter 3, section 3.3.8, 3.3.9, and 3.3.10.

4.3.6 Cloning and construction of CrDL7H in the yeast expression vector

Total RNA from the first pair of *C. roseus* leaves was isolated using *TRIzol*® reagent (Invitrogen). The cDNA synthesis was performed using AMV-Reverse Transcriptase (Promega) with an oligo d(T) primer, according to the manufacturer's instruction. This cDNA pool was used to clone CrDL7H (Genbank™ accession no KF415115) and a NADPH-cytochrome P450 reductase (CrCPR) from *C. roseus* (Genbank™ accession no X69791) and cloned into the pESC-Leu2d yeast dual expression vector (Stratagene, previously described in Ro, Ehlting, and Douglas, 2002; Ro *et al*., 2008). This yeast expression vector was modified by deleting some regions in the promoter of *Leu2d* allele that causes *S. cerevisiae* to produce higher plasmid copy number in compensating the weakened *Leu2* expression. Coding region of CrCPR was

amplified by PCR using High-Fidelity Phusion DNA polymerase (New England Biolabs) using a forward primer (5'- GCGACGGAGTTGGGATTTTAT -3') and a reverse primer (5'- TTGAAAACATCTGGAGGGGTG-3'). The PCR fragment of CrCPR was inserted into pGEMT-Easy vector (Promega), digested with *NotI* and inserted into the corresponding site behind the Gal-10 promoter in the correct orientation. The full-length open reading frame (ORF) for CrDL7H was first cloned into pGEMT-Easy vector (Promega) using a forward primer, 5'-TACAGCGGGCCCAGGATGGAATTGAACTTCAAATC-3' containing *ApaI* site (underlined) and a reverse primer, 5'-GCGCGTCGACGAGTTTGTGCAGAATCAAATGA-3' containing *SalI* site (underlined), excised from the pGEMT-Easy vector and ligated into the corresponding site behind the Gal-1 promoter. Its sequence was submitted to the P450 nomenclature committee (D.R. Nelson) for the assignment of a P450 family name.

4.3.7 Yeast strain growth for in vivo activity

S. cerevisiae MKP-0 was used as the host strain for the CrDL7H expression in pESC-Leu2d. The corresponding empty vector with only CrCPR in the yeast expression vector was also introduced into this yeast strain via the standard polyethylene glycol-lithium acetate procedure (Gietz *et al*., 1992). *S. cerevisiae* harboring the plasmid were grown in selective SD-Leu dropout medium with 2% glucose as a carbon source and protein expression was induced by SG-Leu dropout medium with 2% galactose as a sole carbon source as previously described (Ro, Ehlting, and Douglas, 2002). For *in vivo* assay in yeast, 50 mL of late-log phase yeast cells was harvested at 1,000 g after Gal

induction and resuspended in 25 mL of Tris-EDTA buffer (pH 7.5) with 0.2 mM 7-deoxyloganic acid. The yeast culture was incubated with 7-deoxyloganic acid at 28ºC for 24 h in a 150-rpm shaker, the cells were removed by centrifugation, and the medium was freeze-dried overnight and resuspended in 1 mL of water. The sample was passed through Supelclean™ ENVI-18™ column (Supelco, Inc) equilibrated with 2 mL of acetonitrile followed by 2 mL of water. The sample was eluted with 200 µL of methanol and analyzed by UPLC-MS.

4.3.8 Enzyme assays and substrate specificity

Microsomal preparation from yeast was performed using the high density procedure as described (Pompon *et al.,* 1996). The cell pellets were resuspended in TEK buffer (0.1 M KCl in TE [50 mM Tris-HCl, pH 7.45, 1 mM EDTA]) to a concentration of 0.5 g wet cells per mL followed by incubation at room temperature for 5 min. The cells were recovered by centrifugation and resuspended in the minimum volume of TES B (0.6 M sorbitol in TE) [1 mL per each 100 mL of initial culture]. Glass beads (diameter 0.45/0.5 mm) were gently added until skimming (as exactly as possible) the top of the cell suspension and the cell walls were disrupted mechanically by hand shaking with up and down movements at two movements per second in a cold room for 10 min at 30s intervals separated by 30s intervals on ice. The crude extracts and the glass beads were washed with 5 mL of TES B for three times. The pooled supernatants were placed in new tubes and centrifuged at 4°C for 10 min at 15,000 rpm. After the pellet containing intact cells, nuclei, and disrupted mitochondria were discarded, the precipitation of microsomes was performed by adding NaCl to supernatant to a final concentration of

0.15 M and PEG-4000 to a final concentration of 0.1 g/mL and incubated on ice for at least 15 min. The microsomal fractions were recovered by centrifugation at 4°C for 30 min at 15,000 rpm and the pellets were resuspended in TEG (20% glycerol in TE) at 20-40 mg per mL. The total microsomal protein content was determined using Bradford assay (1976).

For *in vitro* enzyme assay, the incubation mixture (0.1 mL) contained 460 µg of microsomal protein, 1 mM NADPH, 4 mM dithiothreitol. The reaction was initiated by addition of 138 µM 7-deoxyloganic acid to the mixture and incubation for 120 min at 30ºC. For substrate specificity, iridoids tested were 0.1 mM of 7-deoxyloganin, 7-deoxyloganetic acid, 10-deoxygeniposide, and loganin. Assays with boiled microsomes or lacking NADPH served as negative controls. The reaction was terminated by chilling the reaction mixture and water was added to total volume of 1 mL. The sample then was introduced to Supelclean^{TM} ENVI-18^{TM} column (Supelco, Inc) as described above.

Enzyme assays (0.2 mL reaction volume) contained 2.3 mg of CrDL7H enriched microsomes, 4 mM dithiothreitol with serial dilutions of 7-deoxyloganic acid (27.7µm, 55.5µm, 83.3µm, 138 µM, 277µM, 555µM, 694µM, 972µM, 1.51 mM) at 1 mM of NADPH and those for NADPH at 10µM, 15µM, 20µM, 25µM, 30µM, 40µM, 80µM, 100 µM, 120 µM using 0.2 mM of 7-deoxyloganic acid were prepared for kinetic analysis. The apparent Michaelis-Menten constant (K_m) and maximal reaction velocity (V_{max}) for 7-deoxyloganic acid and NADPH was determined by Lineweaver –Burke plots using GraphPad.

4.3.9 Metabolite analysis by liquid chromatography-mass spectrometry

An aliquot of 200 µL of each methanolic extract of a single leaf was filtered through 0.22 µm PALL filter (VWR) before injection to ultra-performance liquid chromatography – mass spectrometry (UPLC-MS) (Waters). The metabolites were separated by using an Aquity UPLC BEH C_{18} with a particle size of 1.7 µm and column dimensions of 1.0 x 50 mm. Samples were maintained at 4°C and 5 µL injection were applied into the column with detection by photodiode arrays and MS. The solvent systems for iridoid analysis were as follows: solvent A, 0.1% formic acid and solvent B, 100% acetonitrile at 0.3 mL/min all times. The following linear elution gradient was used: 0-0.5 min 99% A, 1% B; 0.5-5.0 min 92% A, 8% B; 5.0-6.5 min 70% A, 30% B; 6.5-7.2 min 50% A, 50% B; 7.2-7.5 min 70% A, 30% B, 7.5-8.0 min 92% A 8% B; 8.0 min 99% A, 1% B. All iridoid reference standards were also analyzed by this method and standard curve was generated to quantify the levels of each iridoid in the extracts. The mass spectrometer was operated with a capillary voltage of 3.0 kV, cone voltage of 30 V, cone gas flow of 1L/h, desolvation gas flow of 600 L/h, desolvation temperature of 350°C, and a source temperature of 150 °C. All iridoids were detected at 240 nm with secologanin (positive ion mode, m/z = 389, RT= 3.88 min), 7-deoxyloganic acid (negative ion mode, m/z = 359, RT= 4.06 min), 7-deoxyloganin (positive ion mode, m/z = 375, RT = 5.27 min), 7-deoxyloganetic acid (negative ion mode, m/z= 196, RT = 4.68 min), loganin (positive ion mode, m/z = 391, RT= 3.10 min), and loganic acid (negative ion mode, m/z = 375, RT= 1.90 min). MIAs were analyzed in the same column condition and detection method previously described (Roepke *et al*., 2010).

4.3.10 Real-time quantitative RT-PCR (qRT-PCR)

The expression level of RPP0C (gene encoding 60s acidic ribosomal protein P0-C) was estimated as a housekeeping gene with the following RPP0C specific primers: forward (5'-TCTTAGTTGGAATGTTCAGCACCTG-3') and reverse (5'-CAAGGTTGGAGCCCCTGCTCGTGTT- 3'). Two µL of cDNA was used as a PCR template. The real-time qRT-PCR was performed as follows: 95°C for 3 min, 30 cycles of 95°C for 10 sec, 50°C for 20 sec, and 72°C for 30 sec. Real-time qRT-PCR was performed on an CFX96 real-time SYBR system (Bio-Rad) using detection of SYBR green following Bio-Rad protocol. Primers used for measuring the expression of CrDL7H are a forward primer (5'-CGATGATCTTGTACGAGGTT-3') and a reverse primer (5'-GATGAAGTAGGATTTGTGGC -3').

4.3.11 Phylogenetic analysis

The amino acid alignments were performed using ClustalW (Thompson, Higgins, and Gibson, 1994). The neighbor-joining phylogeny was generated with MEGA 5.1 with bootstrap analysis of 10000 replicates (Tamura *et al*., 2011).

4.4 Results

4.4.1 Bioinformatic-guided screening of the PhytoMetaSyn database provides 3 CYP candidate genes to be tested for the 7-deoxyloganic acid 7-hydroxylase function

Screening of transcript databases assembled and annotated for the PhytoMetaSyn project (http://www.phytometasyn.ca/) (Facchini *et al*., 2012) from various MIA producing Apocynaceae plants (*Vinca minor, Rauwolfia serpentina, Tabernaemontana*

elegans, Amsonia hubrichtii), from *Cinchona ledgeriana* (Rubiaceae) (O'Connor and Maresh, 2006), and from secologanin producing *Lonicera japonica* (Caprifoliaceae) (Kawai, Kuroyanagi, and Ueno, 1988) identified several putative known iridoid biosynthetic genes from *C. roseus* (Table 4.S1). Remarkably all six species contain candidate genes for each of the known iridoid biosynthetic genes including the recently characterized geraniol synthase (GES) (Simkin *et al*., 2013), G10H, 10-HGO, IRS, LAMT and SLS (Table 4.S2). These results triggered a search for cytochrome P450s with 80-90% amino acid sequence identity to identify candidate genes that might be involved in iridoid biosynthesis. The screening of initial hits against non-secologanin producing species (*Abies balsamea, Saponaria vaccaria, Cannabis sativa, Lactuca sativa, Papaver bracteatum*) eliminated genes likely to be involved in other pathways, thus yielding *CrCYP72A224* as well as 2 other CYPs that were selected for further study (Table 4.S1).

4.4.2 VIGS of CrCYP72A224 modifies the iridoid metabolite profile of silenced Catharanthus plants

In order to test their possible roles in secologanin biosynthesis, each of the 3 genes were submitted to a modified virus-induced gene silencing (VIGS) method (De Luca *et al*., 2012b) based on the successful identification of new genes in iridoid (Geu-Flores *et al*., 2012) and MIA (Liscombe and O'Connor, 2011) biosynthesis in *C. roseus* using a well-known pTRV vector system (Burch-Smith *et al*., 2004) that has been successfully applied in several plant species (De Luca *et al*., 2012a). As described previously (De Luca *et al*., 2012b), additional experiments to silence the phytoene

desaturase (PDS) gene were also performed since the visible photobleached phenotype obtained could be used to decide when it was time to perform transcript and metabolite profiling in plants suppressed for each of the 3 CYP genes being tested. The effect of transcriptional down-regulation of *CrCYP72A224* on the iridoid profiles of silenced *Catharanthus* tissues were monitored by UPLC-MS. Remarkably, the silencing resulted in dramatic drops of secologanin in *CrCYP72A224-vigs* silenced plants (Figure 4.2, *CrDL7H -vigs-1* and *vigs-2*) compared to the levels of secologanin found in plants treated with empty vector (Figure 4.2, EV). The silencing of *CrCYP72A224* also resulted in the appearance of 7-deoxyloganic acid and its peak was confirmed by UV spectrum and mass of 7-deoxyloganic acid standard (Figures 4.2, 4.S1A and B). Since the loss of secologanin resulted in the accumulation of 7-deoxyloganic acid (Figure 4.2, *CrDL7H -vigs-1* and *vigs-2*) the putative substrate of the *CrDL7H* (Figure 4.1) reaction, we decided to use this nomenclature for *CrCYP72A224*.

Transcript analysis of *CrDL7H* monitored by real time PCR revealed that they declined by about 75% (P value of <0.05) in leaves infiltrated with *A. tumefaciens* harboring pTRV2-CrDL7H compared with those exposed to bacteria containing pTRV2-EV (Figure 4.3A). More detailed metabolite analyses of *CrDL7H* -silenced tissues showed that secologanin levels declined by at least 70 % compared with non-inoculated (WT), EV- or Mock-inoculated plants, respectively (Figure 4.3B). Remarkably *CrDL7H*-silenced tissues accumulated over 4 mg/gFW of 7-deoxyloganic acid (Figure 4.3B) compared to *Catharanthus* control, empty vector and mock inoculated tissues. MIA analyses of *CrDL7H*-silenced tissues showed that they responded by a > 60 % decline of

catharanthine and > 30 % decline of vindoline accumulation compared to *Catharanthus* control, empty vector, and mock inoculated tissues (Figure 4.3C).

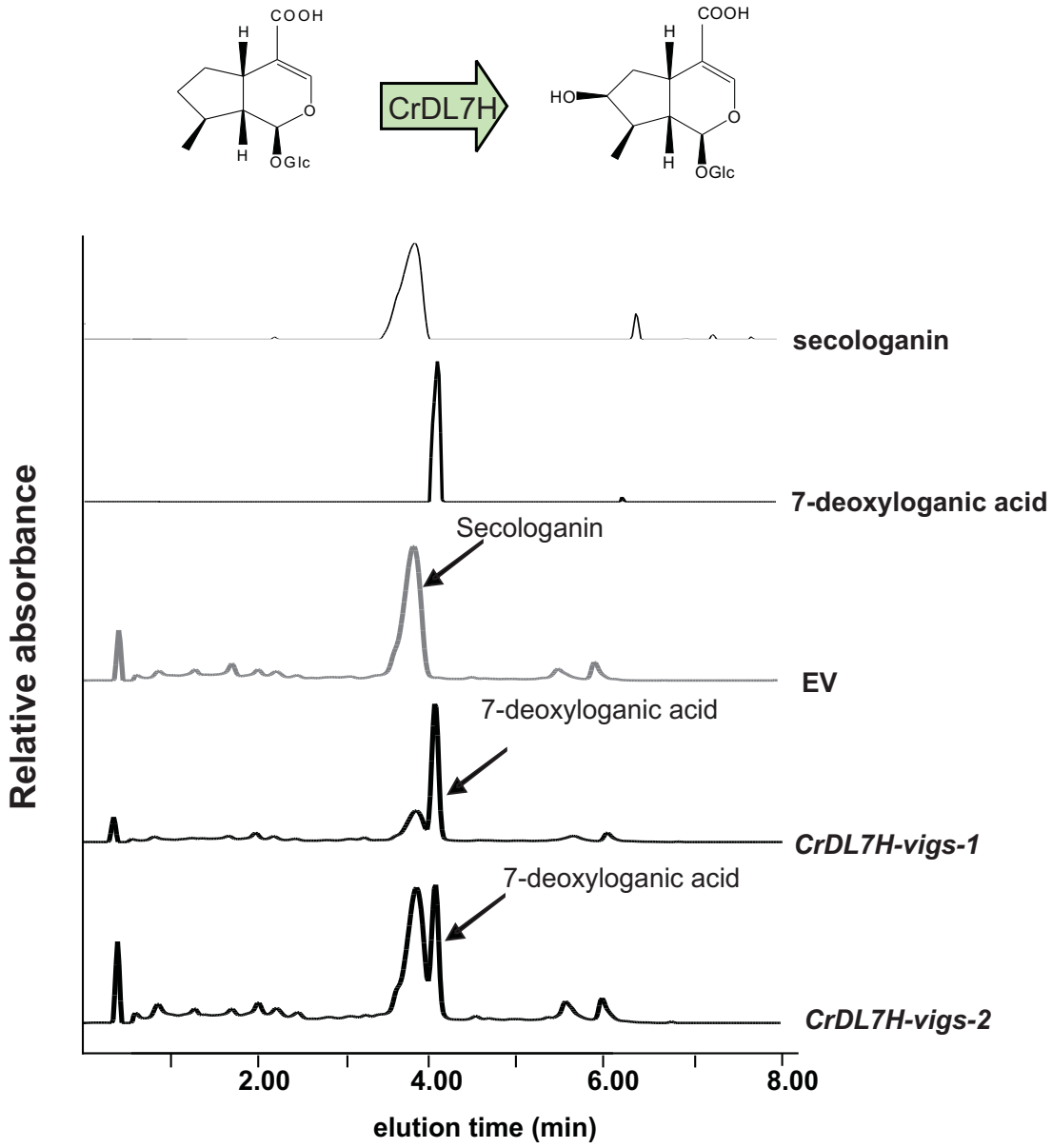


Figure 4.2 Selected LC-MS chromatograms of two VIGS-treated plants (*DL7H-vigs-1* and *DL7H-vigs-2*). When CrDL7H is silenced, the plants accumulate 7-deoxyloganic acid and decreased level of secologanin compared with empty vector (EV).

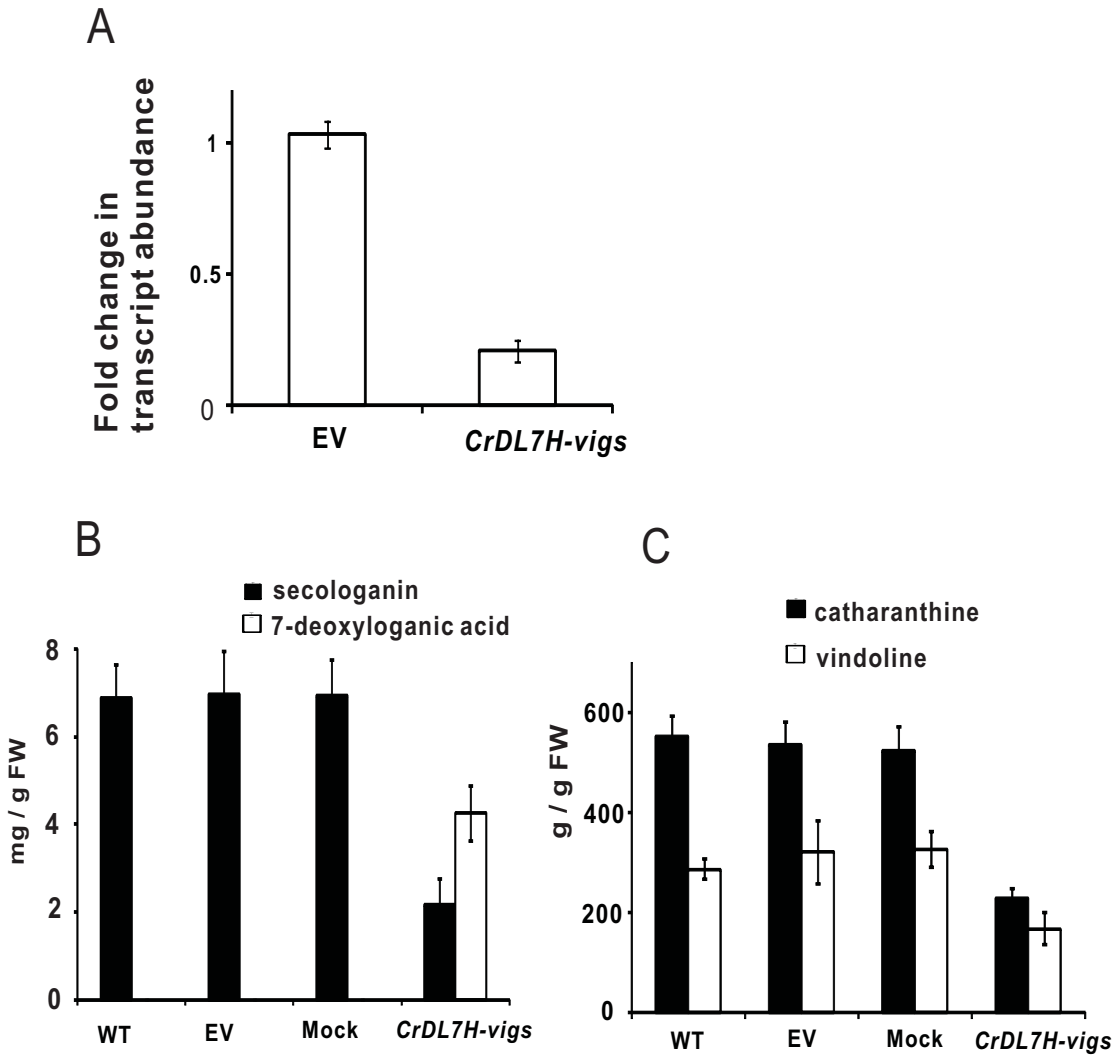


Figure 4.3 Effects of silencing CrDL7H in *C. roseus*. (A) *CrDL7H* transcript levels in plants infiltrated with *A. tumefaciens* harboring pTRV2-CrDL7H was determined by qRT-PCR to measure the fold change of transcript levels from individual plants of silenced gene relative to mock infected plants. Data represents the mean ± standard error of at least eight biological replicates performed on cDNA prepared for each individual plants infiltrated with *A. tumefaciens* harboring each construct. **(B)** Effect of suppressing *CrDL7H* transcript levels on the iridoid contents and **(C)** major alkaloids, catharanthine and vindoline of *C. roseus*. Leaves from plants infiltrated with *A. tumefaciens* containing each construct. Data is presented in mean ± standard error of at least eight biological replicates. Determination of catharanthine and vindoline levels from wild-type, EV and VIGS treated plants showed that the production of MIAs was impaired when secologanin biosynthesis was suppressed by silencing gene involved in the iridoid pathway. Significant differences were considered with P values of <0.05 by Student's *t*-test for the transcript analysis and metabolite contents of EV-infected plants and each silenced lines.

4.4.3 Functional characterization of CrDL7H

To functionally characterize *CrDL7H*, an *in vivo* assay was performed using an engineered dual expression yeast system, pESC-Leu2d containing *Catharanthus roseus* NADPH-cytochrome P450 reductase (*CrCPR*) as a redox partner and the candidate gene that were under the control of a galactose-inducible promoter (Ro, Ehlting, and Douglas, 2002). Supplementation of the induced *CrDL7H* expressing yeast culture with 7-deoxyloganic acid led to the production of a novel more hydrophilic iridoid (Figure 4.S1C) with the same retention time (RT = 1.9 min) and mass (m/z = 375) of authentic loganic acid (Figure 4.S1C), the 7-hydroxy-derivative of 7-deoxyloganic acid (Figure 4.4) while no biotransformation was observed in the yeast cultures transformed with empty pESC-Leu2d expressing only *CrCPR* (Figure 4.4).

To verify the results obtained with the *in vivo* biotransformation of 7-deoxyloganic acid into loganic acid (Figure 4.4), microsomes prepared from *CrDL7H* expressing yeast cultures were assayed for the presence of DL7H activity. In the presence of NADPH, yeast microsomes were able to convert 7-deoxyloganic acid to a product (Figure 4.S1D) with the same retention time (RT = 1.9 min), UV spectra and mass (m/z 375) as authentic loganic acid (Figure 4.S1C and Figure 4.5). Assays lacking NADPH cofactor (Figure 4.5) or those incubated in the presence of boiled *CrDL7H*-enriched microsomes served as negative controls. While the NADPH could be replaced with NADH, this co-substrate was only 20 % as active. Other iridoid substrates (7-deoxyloganin, 7-deoxyloganetic acid, 10-deoxygeniposide, and loganin; Figure 4.S2) tested were not accepted by the recombinant enzyme. Kinetic analyses with yeast microsomes expressing this enzyme (Figure 4.S3) yielded apparent of

K_m (V_{max})= 111.07 ± 14.8 μM (5.5 ± 0.77 μM min^{-1} μg protein) and 29.85 ± 3.28 μM (4.19 ± 0.40 μM min^{-1} $μg^{-1}$ protein) for 7-deoxyloganic acid and NADPH, respectively. The K_m obtained for the enzyme expressed in yeast is similar to that of the enzyme from microsomes of *Lonicera japonica* cell cultures (Katano, 2001).

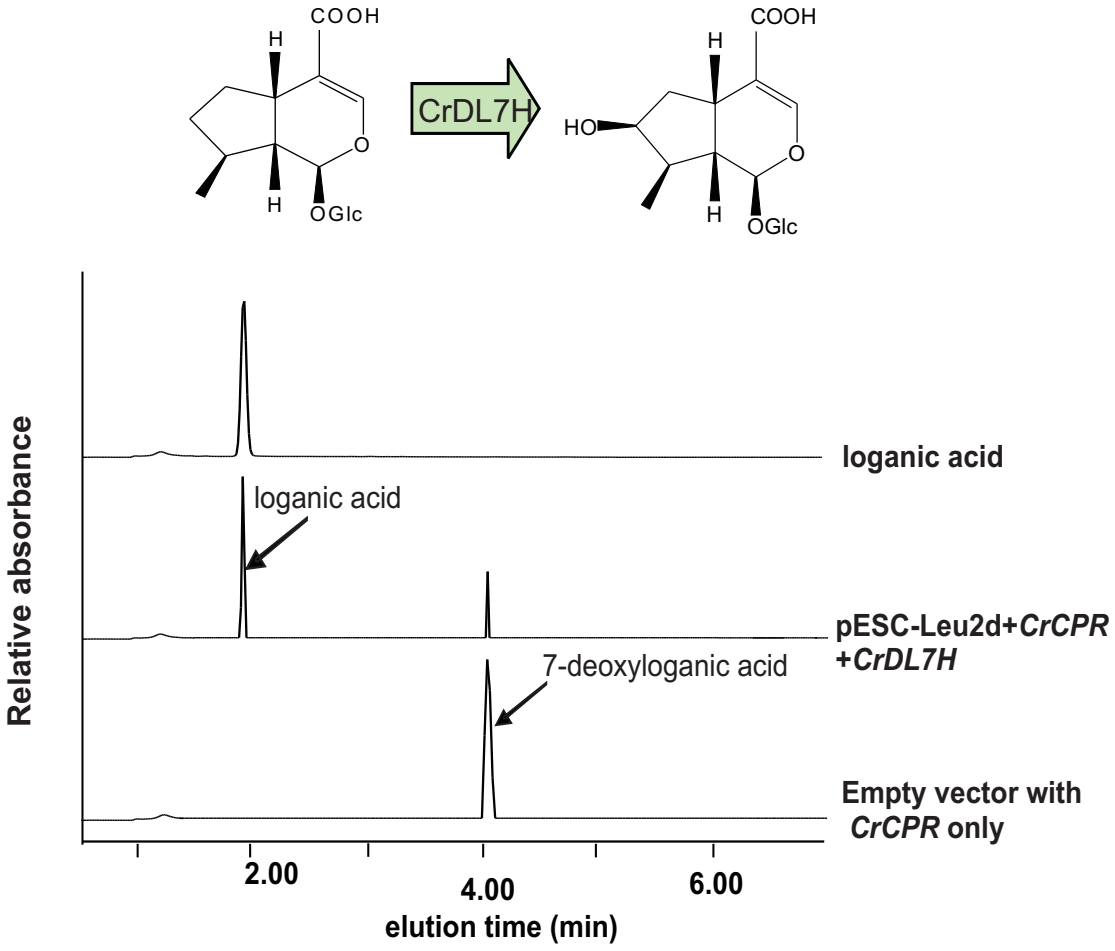


Figure 4.4 ***In vivo*** **characterization of recombinant** ***CrDL7H*****,** with selected ion chromatograms showing conversion of 7-deoxyloganic acid (m/z= 359) to loganic acid (m/z = 375). Production of loganic acid only occurred in the recombinant enzyme *CrDL7H*, but not in the reactions with the empty vector containing *CrCPR* only.

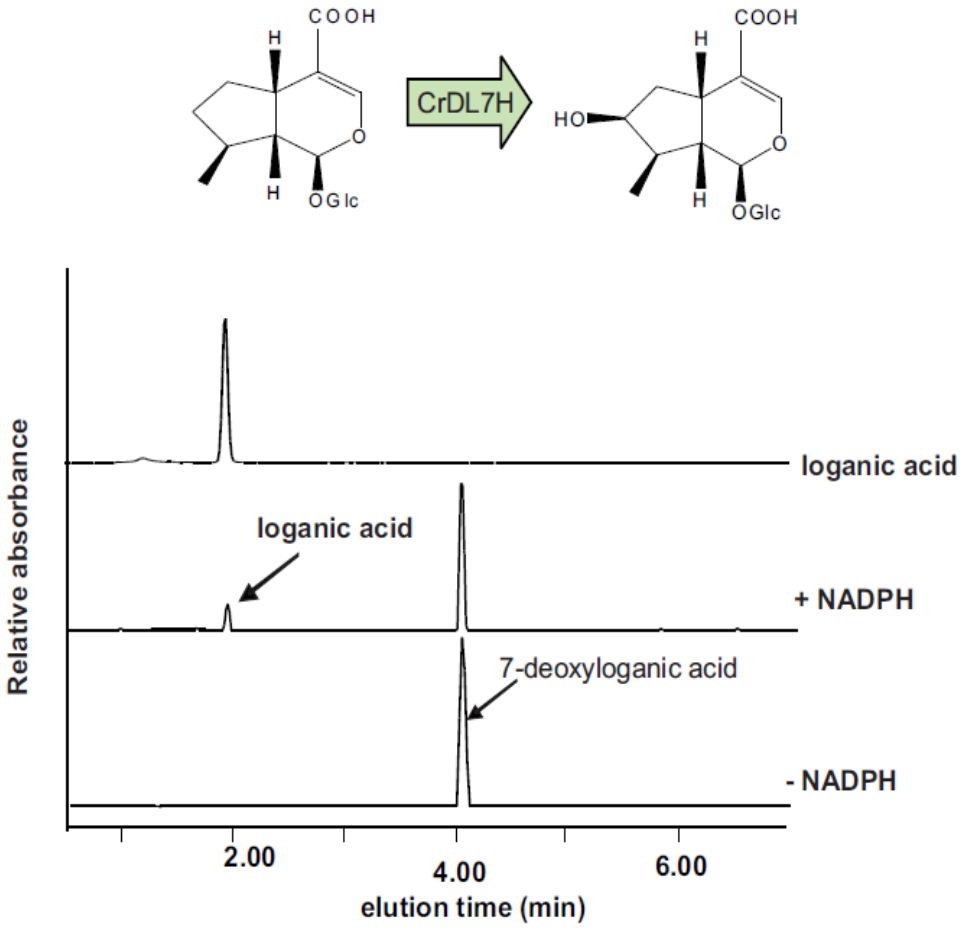


Figure 4.5 Production of loganic acid by yeast microsomal extracts expressing recombinant *CrDL7H*. Loganic acid production only observed in the presence of recombinant enzyme, NADPH, and 7-deoxyloganic acid, but not in the reactions without NADPH.

4.4.4 Preferential expression of CrDL7H in different Catharanthus plant organs

The expression levels of *CrDL7H* in various *C. roseus* organs were examined by real-time quantitative reverse transcriptase-PCR (qRT-PCR) and their relative abundance was compared to the levels of secologanin, catharanthine and vindoline (Figure 4.6A). The transcript level of *CrDL7H* was highest in leaf pair 1 and it declined with leaf age as seen in leaf pairs 2 to 5 (Figure 4.6A). The patterns of expression of *CrDL7H* were very similar to those of *LAMT* and *SLS* (Chapter 3, Figure 3.3) that coordinates with the pattern of secologanin and MIA accumulation within these organs (Figure 4.6A). The levels of *CrDL7H* were consistently higher in actively growing aerial organs as also illustrated when comparing expression in flower buds compared to more mature open flowers (Figure 4.6A). Expression of *CrDL7H* was also detected in stems and roots consistent with the detection of secologanin and/or MIAs in these organs (Figure 4.6A; Chapter 3, Figure 3.3) and with the role of this enzyme in catalyzing the 3rd to last step in secologanin biosynthesis in *Catharanthus* (Figure 4.1).

4.4.5 CrDL7H was not preferentially expressed in epidermal cells

Carborundum abrasion technique has been successfully developed in our laboratory to harvest leaf epidermis enriched RNA from *C. roseus* and to corroborate MIA pathway gene expression in these cells (Murata et al., 2008; De Luca *et al*., 2012b). Isolated epidermis enriched transcripts analysed by qRT-PCR showed that transcripts for loganic acid *O*-methyltransferase (*LAMT*) and secologanin synthase (*SLS*), well known to be preferentially expressed in the leaf epidermis (Guirimand *et al*., 2011; Murata *et al*., 2008), were 2.5- and 5-fold more enriched in these cells than in whole leaf extracts

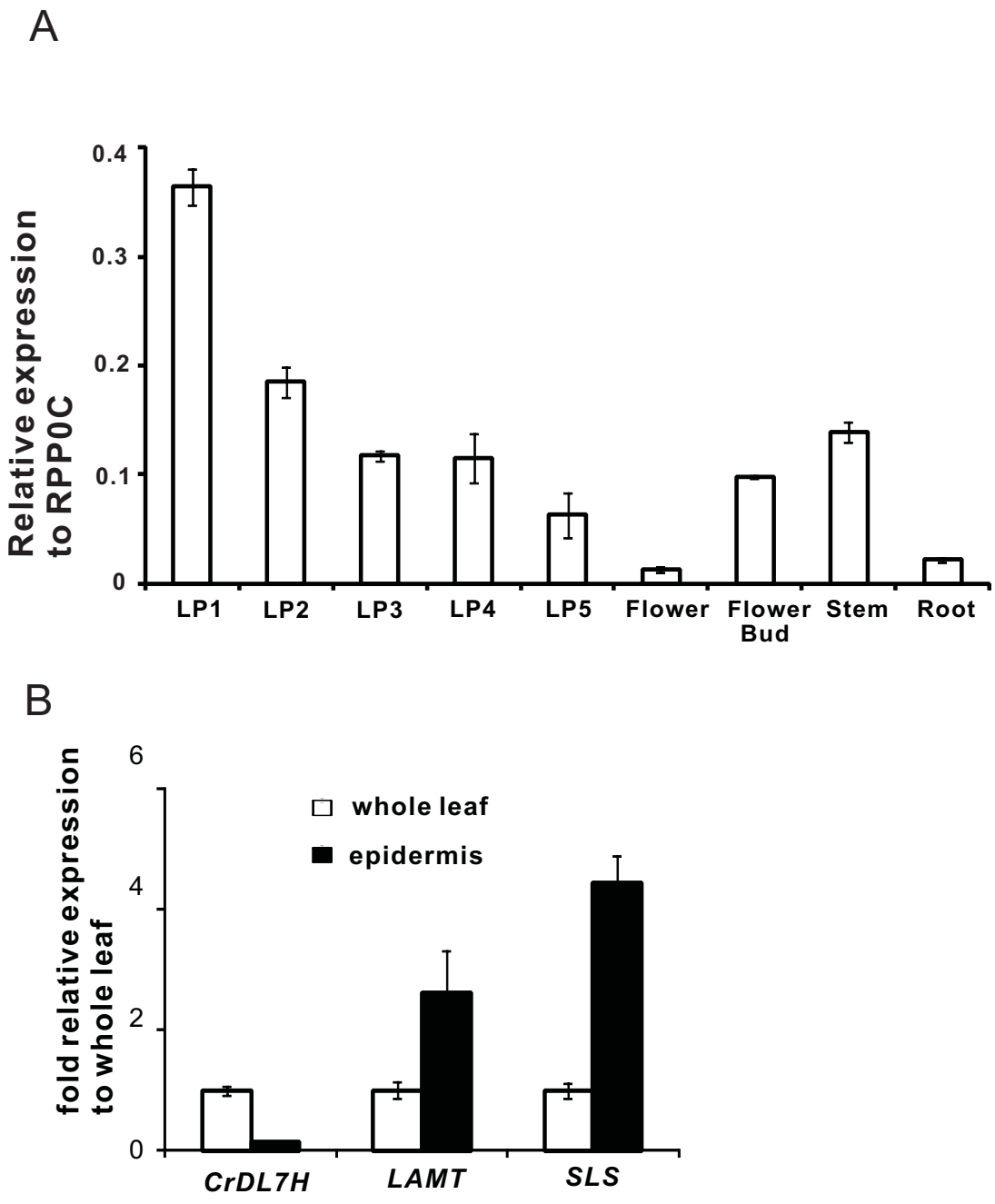


Figure 4.6 Preferential expression of CrDL7H. (A) Expression of *CrDL7H* in various organs of *C. roseus* was measured by quantitative real-time RT-PCR performed in three replicates. Data are represented as a relative expression to RPP0C (gene encoding 60S acidic ribosomal protein P0-C). **(B)** Transcript levels of the three genes involved in secologanin biosynthesis in epidermis relative to whole leaf of *C. roseus*. Unlike *SLS* and *LAMT*, *CrDL7H* transcript is not highly abundant in the epidermis. SLS-secologanin synthase, LAMT-loganic acid *O*-methyltransferase, CrDL7H-7-deoxyloganic acid 7-hydroxylase. *SLS* and *LAMT* expression were measured previously in Chapter 3 (Figure 3.4) for comparison with *CrDL7H* expression in this figure.

(Chapter 3, Figure 3.4, 4.6B). In contrast, *CrDL7H* transcripts were over 7-fold more enriched in whole leaves than in the leaf epidermis (Figure 4.6B). Together these results suggested that *CrDL7H* is preferentially expressed within *Catharanthus* leaves rather than in leaf epidermal cells where the last 2 steps in secologanin assembly are expressed.

4.4.6 Phylogenetic analysis of CYPs that are related to CrDL7H

Since CrDL7H belongs to the CYP72A subfamily, it was designated CYP72A224. Remarkably, CYP72A224 shows the highest degree of similarity to a putative secologanin synthase-like gene from *Camptotheca acuminata* (GenBank accession #: HQ605982 with 56% identity), while it shares 54% amino acid identity with *C. roseus* SLS (CYP72A1) that has been functionally characterized (Irmler *et al*., 2000). All species that accumulate secologanin (*Lonicera japonica*) and/or MIAs (*Cinchona ledgeriana*, *Vinca minor, Tabernaemontana elegans, Rauwolfia serpentina, Amsonia hubrichtii*) appear to contain CrDL7H-like candidate genes as well as CYP candidates for CYP72A1 and G10H from the CYP76B subfamily. A phylogenetic tree (Figure 4.7) illustrates the relationship of CYPs involved in iridoid biosynthesis to 4 functionally characterized CYP72A genes involved in triterpene biosynthesis (Seki *et al*., 2011; Fukushima *et al*., 2013).

CrDL7H also shows 49% and 48 % amino acid identities to CYP72A154 from *Glycyrrhiza uralensis* and CYP72A63 from *Medicago truncatula* involved in C-30 oxidation in the glycyrrhizin and β-amyrin biosynthesis (Seki *et al*., 2011), respectively. In addition CYP72A61v2 and CYP72A68v2 that are 49% identical to CrDL7H from *Medicago truncatula* performed unusual oxidations of triterpenoids when they were

expressed in transgenic yeast (Fukushima et al., 2013). The phylogenetic analysis (Figure 4.7) clearly shows that both the CrDL7H and SLS genes from the 6 species grouped separately from the other members of the CYP72A family involved in triterpene biosynthesis as well as from 4 other *C. roseus* CYPs with known biochemical functions (flavonoid 3',5'-hydroxylase (CYP75A8) (Kaltenbach *et al*., 1999), cinnamate 4-hydroxylase, C4H (CYP73A4) (Hotze, Schröder, and Schröder, 1995), tabersonine 16-hydroxylase, (T16H; CYP71D12) (Schröder *et al*., 1999), and tabersonine 19-hydroxylase, (T19H; CYP71BJ1) (Giddings *et al*., 2011).

4.5 Discussion

4.5.1 Combined use of transcriptomics, VIGS, and metabolite analyses identifies CrDL7H, a member of the CYP72A subfamily that catalyzes the 3rd to last reaction in secologanin biosynthesis

This study provides a successful bioinformatic approach, using transcriptomic data from seven plant species that produce secologanin (*Lonicera japonica*) and/or MIAs (*Catharanthus roseus, Vinca minor, Tabernaemontana elegans, Rauwolfia serpentina, Amsonia hubrichtii*, and *Cinchona ledgeriana*) to identify candidate cytochrome P450 genes that have novel functions in iridoid biosynthesis. The search identified a secologanin synthase-like CYP72A subfamily member (CYP72A224) in addition to SLS (CYP72A1) that was previously functionally characterized from *C. roseus* cell cultures (Irmler *et al*., 2000). While the CYP72A subfamily are widely distributed among plants including the Arabidopsis and rice model systems

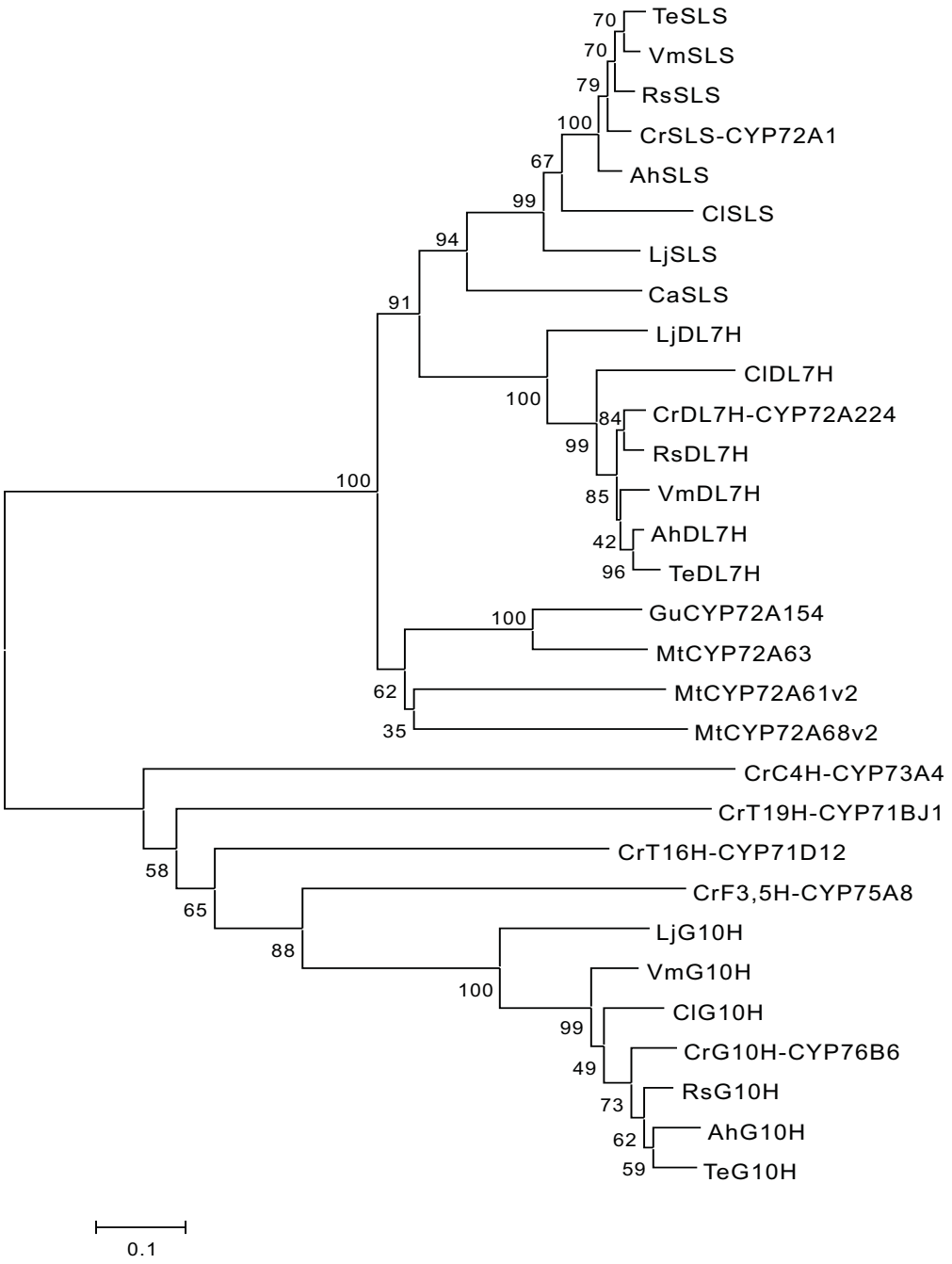


Figure 4.7 Phylogenetic relationships among relevant cytochrome P450s to CYP72A proteins. A phylogenetic tree was generated using the neighbor-joining algorithm based on a comparison of amino acid sequences using the ClustalW program. The numbers on each branch represent bootstrap values for 10000 replicates. The bar indicates 0.1 substitutions per position in the sequence. Abbreviations: Cr-*Catharanthus roseus*, Vm-*Vinca minor*, Rs-*Rauwolfia serpentina*, Te-*Tabernaemontana elegans*, Ah-*Amsonia hubrichtii*, Cl-*Cinchona ledgeriana*, Lj-*Lonicera japonica*, Ca-*Camptotheca acuminata*, Gu-*Glycyrrhiza uralensis*, Mt-*Medicago truncatula*, G10H-geraniol 10-hydroxylase, SLS-secologanin synthase, DL7H-7-deoxyloganic acid 7-hydroxylase, C4H-cinnamate 4-hydroxylase, F3,5H-flavonoid 3',5'-hydroxylase, T16H-tabersonine 16-hydroxylase, T19H-tabersonine 19-hydroxylase. The GenBank accession numbers of these CYPs were listed in Table 4.S3.

(http://drnelson.uthsc.edu/biblioD.html) (Nelson *et al*., 2008), four CYP72A subfamily genes that have been functionally characterized are involved in triterpene biosynthesis (Seki *et al*., 2011; Fukushima *et al*., 2013). In another example, 4 clustered P450s from the same CYP71C family are involved in 2,4-dihydro-1,4-benzoxazin-3-one (DIBOA) and 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) biosynthesis in the Gramineae (Frey *et al*., 1997). Interestingly, these clustered genes also share an overall amino acid identity of 45 to 60% and show strict substrate-specificities like CYP72A1 and CYP72A224 that are involved in separate steps in secologanin biosynthesis.

While a few *C. roseus* P450 genes have been cloned by traditional screening methods combined with *in vitro* enzyme assays of recombinant proteins for determining substrate specificities (Mizutani and Ohta, 2010), several recent efforts have demonstrated the effectiveness of VIGS for transiently down regulating pathway steps in a growing number of plant species (De Luca *et al*., 2012a) and for identifying appropriate candidate genes as in the present study. The silencing of *CrDL7H* (*CYP72A224*) in *C. roseus* decreased both secologanin and MIA levels in infected plants and led to the appearance of 7-deoxyloganic acid that is not usually detected in plant tissues (Figure 4.2). In addition, *CrDL7H* expressing yeast converts 7-deoxyloganic acid efficiently to loganic acid (Figure 4.4) and *in vitro* enzyme assays with *CrDL7H*-containing microsomes are shown to require NADPH to convert this iridoid to loganic acid, while control yeast cells or microsomes harboring the empty expression vector were inactive (Figure 4.5). The kinetic analysis of *CrDL7H* showed its high affinity for 7-deoxyloganic acid and its inability to oxidize 7-deoxyloganetic acid, 7-deoxyloganin or 10-deoxygeniposide strongly suggests that secologanin biosynthesis preferentially involves

hydroxylation followed by methylation of the carboxyl group (Figure 4.1) rather than the alternative route. This route is also supported by the high substrate specificity of recombinant *Catharanthus LAMT* (Murata *et al*., 2008) for loganic acid over 7-deoxyloganic acid. The accumulation of 7-deoxyloganic acid but not 7-deoxyloganin (Figure 4.2) in *CrDL7H-vigs* plants is also consistent with hydroxylation occurring before methylation in the biosynthesis of secologanin in *C. roseus*. Previous studies using cell culture from *Lonicera japonica* and *C. roseus* (Irmler *et al*., 2000; Katano *et al*., 2001) indicated a more flexible order of hydroxylation and methylation (Figure 4.1) while others (Uesato *et al*., 1986) suggested that hydroxylation occurred at the stage of 7-deoxyloganic acid based on feeding various labeled iridoid and secoiridoid glucosides to *Galium* species. The earlier study reporting the ability of *L. japonica* cell cultures to convert 7-deoxyloganin to loganin (Yamamoto *et al*., 1999) deserves further investigation by cloning the relevant methyltransferase for functional expression and biochemical analysis.

4.5.2 Early steps of iridoid biosynthesis in C. roseus are preferentially expressed in IPAP cells

The gradual decrease of *CrDL7H* transcripts with leaf development (Figure 4.6; leaf pairs 1 to 5) reflects the results obtained with the decline in LAMT enzyme activity (Murata *et al*., 2008) and the pattern is consistent with the levels of secologanin and catharanthine accumulation in various organs. While expression of *LAMT* and *SLS* is restricted to leaf epidermal cells (Irmler et al., 2000; Murata *et al*., 2008), qRT-PCR analysis of epidermal enriched cDNA showed that *CrDL7H* transcripts are not highly

expressed in epidermal cells (Figure 4.6B). The biosynthesis of monoterpenoid precursors in aerial organs of *C. roseus* has been shown to involve at least two different cell types, epidermal cells and internal phloem-associated parenchyma (IPAP) cells where geraniol synthase, G10H, iridoid synthase, and CrUGT8 are expressed (Simkin *et al*., 2013; Burlat *et al*., 2004; Geu-Flores *et al*., 2012; Chapter 3). These data suggest the importance of CrDL7H characterization to determine if 7-deoxyloganic acid or loganic acid is transported to the epidermis. Although the accumulation of the gene transcript of CrDL7H in non-epidermal cells, such as IPAP cells need to be verified by *in-situ* hybridization studies, these results may suggest the intercellular translocation of intermediate loganic acid between IPAP cells and epidermis (Figure 4.8). The preferential expression of CrDL7H in non-epidermal cells was also confirmed by its absence in the epidermis cDNA library (Murata *et al*., 2006) (Figure 4.S2). The data presented here suggests that the search for loganic acid transporter would be an interesting subject to elucidate the transport mechanism of intermediate iridoid between IPAP cells and epidermis.

4.6 Conclusions

In conclusion, we report the functional characterization of CYP72A224 that hydroxylates 7-deoxyloganic acid using virus-induced gene silencing approach followed by recombinant expression in yeast. The identification of CrDL7H not only provides insight into the detailed reactions in iridoid biosynthesis pathway and their localization in IPAP cells and epidermis, it also reveals the potential of metabolic engineering using plants or microbial host for the production of valuable monoterpenoid indole alkaloids.

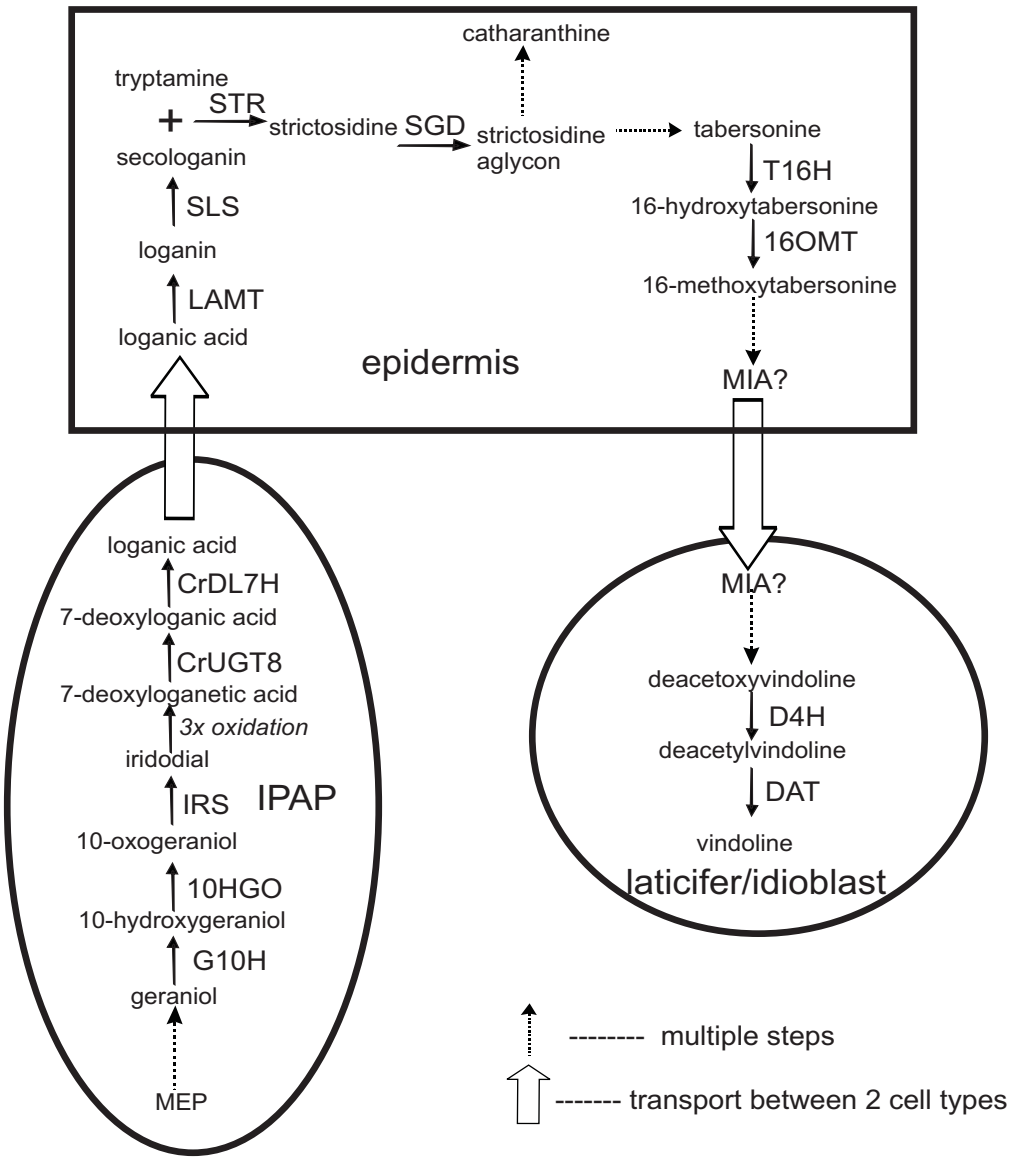


Figure 4.8 Spatial model of iridoid and MIA biosynthesis that involves IPAP, epidermis, and idioblast/laticifer cells. The MEP pathway and iridoid biosynthesis to loganic acid take place in IPAP cells while the last two steps, LAMT and SLS reactions occur in the leaf epidermis. The model illustrates the localization of 7-deoxyloganic acid 7-hydroxylase in IPAP cells. The assembly of strictosidine from secologanin and tryptamine by STR also occurs in leaf epidermis. Another translocation of unknown tabersonine-derived MIA remains to be elucidated as T16H and 16OMT are localized in leaf epidermis and the last two steps of vindoline biosynthesis, D4H and DAT are preferentially expressed in idioblast/laticifer cells, while catharanthine biosynthetic genes are predicted to be localized in epidermis. Abbreviations: MEP-methylerythritol phosphate; IPAP-internal phloem associated parenchyma; G10H-geraniol 10-hydroxylase; 10HGO-10-hydroxygeraniol oxidoreductase; IRS-iridoid synthase; CrUGT8- 7-deoxyloganetic acid glucosyltransferase; CrDL7H- 7-deoxyloganic acid 7-hydroxylase; LAMT-loganic acid *O*-methyltransferase; SLS-secologanin synthase; STR-strictosidine synthase; SGD-strictosidine β-D-glucosidase; T16H-tabersonine 16-hydroxylase; 16OMT- 16-hydroxytabersonine *O*-methyltransferase; D4H-deacetoxyvindoline 4-hydroxylase; DAT-deacetylvindoline 4-acetyltransferase.

The present study reveals that while the last two steps in secologanin biosynthesis occur within the leaf epidermis, early steps of iridoid biosynthesis from MEP pathway to hydroxylation occur in IPAP cells. Functional analysis of another P450 candidate involved in oxidation steps from iridodial to 7-deoxyloganetic acid in *C. roseus* is in progress as it would also uncover rate-limiting steps in iridoid biosynthesis and its regulation.

4.7 Acknowledgements

This work was funded by National Sciences and Engineering Research Council of Canada to V.D.L, who holds the Canada Research Chair in Plant Biotechnology. We also acknowledge funding by Genome Canada, Genome Alberta, Genome Prairie, Genome Quebec, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners. We thank Dr. Peter Facchini for providing the VIGS vectors and Dr. Dae Kyun Ro for providing the yeast expression system. We recognize the skilled technical work of next-generation sequencing personnel at the McGill University-Genome Québec-Innovation Centre. We are grateful to Dr. Christoph Sensen, Mei Xiao, and Ye Zhang for their dedicated bioinformatic support and large scale gene annotation efforts that helped in the identification of genes from the PhytoMetaSyn website.

General Conclusions

Three distinctive studies described in this thesis highlight the elucidation of the biosynthetic machinery that *Catharanthus roseus* utilizes to produce a wide range of complex monoterpene indole alkaloids (MIAs). The first study illustrated in Chapter 2 reveals the spatial separation of monomers, catharanthine and vindoline that explains the low level of dimers found in plants. This chapter also demonstrates the biological role of secretion of catharanthine to the leaf surface for plant defense. The site of accumulation of catharanthine and vindoline confirms the complex organization of how each biosynthetic step is compartmented in multiple cell types, namely the internal phloem-associated parenchyma (IPAP), the epidermis, and the laticifer/idioblast cells. The results presented in Chapter 2 confirm a biological model for the sequestration of catharanthine and vindoline into separate sites, and support the localization of the last two steps of vindoline biosynthesis in the idioblast/laticifer cells. This study also highlights the ability of catharanthine to inhibit the growth of fungus on the leaf surface and to act as an important deterrent to herbivores as this discovery may lead to further search for more plants that secrete alkaloids.

In order to complete our understandings of the complexity of plant secondary metabolic pathways in *Catharanthus roseus*, more biosynthetic genes involved in the formation of MIAs still need to be isolated. Various large-scale transcriptomic projects have recently facilitated the discovery of more genes associated with MIA biosynthesis. In this thesis, candidate MIA biosynthetic genes were obtained from recent large-scale sequencing project, PhytoMetaSyn that has included a wide variety of medicinal plants that produce secondary metabolites with high values (http://www.phytometasyn.ca/). The

methodology of searching for novel biosynthetic genes from *C. roseus* based on comparisons among Apocynaceae datasets has been successfully applied for functional characterization of cytochrome P450 monooxygenase involved in iridoid biosynthesis pathway, 7-deoxyloganic acid 7-hydroxylase (DL7H) (Chapter 4). Interests in mining the dataset for novel cytochrome P450s become important as MIA skeletons often undergo oxygenation by cytochrome P450 activity, which may be followed by further modifications that increase the diversity of the MIAs.

While genomics approaches facilitate the identification of several candidate genes, we employ virus-induced gene silencing (VIGS) to narrow down the number of candidate genes involved in the MIA biosynthesis pathway before performing further enzyme characterization. VIGS has become a promising tool for gene characterization studies. In Chapter 3, VIGS confirms the role of CrUGT8 in iridoid biosynthesis pathway and resulted in a decreased level of secologanin and MIAs within the silenced lines. VIGS was also used to suppress two other genes involved in secologanin biosynthesis, LAMT and SLS. Interestingly, the silencing of these two last steps in iridoid pathway resulted in a more than 50% decline in secologanin and accumulation of their respective substrates, loganic acid and loganin. In this case, VIGS offers another advantage of providing substrates of biochemical analysis, especially if the substrate is difficult to obtain by organic synthesis.

VIGS can also be used to screen a large number of candidate genes obtained from databases generated in large-scale transcriptomic projects. In Chapter 4, VIGS was used to screen cytochrome P450 candidates that may be involved in secologanin biosynthesis. The silencing of CrDL7H resulted in reduced secologanin level and appearance of 7-

deoxyloganic acid that is not usually accumulated in *Catharanthus* plants. Functional expression of CrDL7H further verified its biochemical activity showing its preference for 7-deoxyloganic acid over other closely related substrates. The functional characterization of CrDL7H in assembly of secologanin clarifies the uncertainties concerning the precise order of hydroxylation and methylation in iridoid pathway of *C. roseus*.

The cloning and functional characterization of two genes, glucosyltransferase, CrUGT8 and hydroxylase, CrDL7H should also reveal any translocation between IPAP and epidermis cells in order to understand spatial separation between early and later iridoid pathways. Comparison of *CrUGT8* transcripts between whole leaf and epidermis showed that CrUGT8 is preferentially expressed within *Catharanthus* leaves (Chapter 3). This data is also validated by their absence in the epidermis enriched cDNA library database, and more importantly by *in situ* hybridization studies that localize the preferential expression of CrUGT8 in IPAP cells. The similar analysis using carborundum abrasion method with CrDL7H suggests that this hydroxylase is also preferentially within the *Catharanthus* leaves, although *in situ* hybridization studies need to be performed to determine its exact location.

Future Directions

The involvement of at least three cell types, namely IPAP, epidermis, and idioblast/laticifers indicates the complexities in MIA biosynthesis. Characterization of CrUGT8 and CrDL7H revealed that the translocation of loganic acid is necessary as the methyltransferase (LAMT) that converts this intermediate to loganin has been shown to be localized in epidermis. The future works may involve an attempt to characterize the

transporter between the IPAP cells and epidermis along with its mechanism to investigate the detailed regulation of the iridoid pathway. Since catharanthine is secreted to the waxy surface of the leaf, an identification of a transporter also becomes another interesting subject to investigate and further reveal the mechanism of its secretion.

As VIGS has been shown to be an effective tool in gene function studies, this approach can be used to find the remaining genes involved in the MIA biosythesis pathway, such as the iridodial oxidoreductase, post-strictosidine reactions and towards the assembly of catharanthine. Narrowing down the candidate genes for catharanthine pathway can be performed by comparing the whole leaf and epidermis enriched cDNA library databases as the genes involved in catharanthine pathway are expected to be found in the epidermis following the direction of its secretion to the waxy leaf surface.

The only gene in secologanin biosynthesis pathway that remains to be characterized is an oxidoreductase involved in a three step oxidation, iridodial oxidoreductase to form a carboxyl group of 7-deoxyloganetic acid. Future works may involve the functional characterization of this enzyme in recombinant system to complete our understandings of steps in iridoid biosynthesis pathway in *C. roseus* for metabolic engineering purposes. In addition, the identification of this oxidoreductase would be useful for localization studies although this gene is predicted to be preferentially expressed in IPAP cells as IRS and CrUGT8 have been shown to be localized in this cell type.

While *C. roseus* is the most well-studied system for investigating the MIA biosynthesis pathway, the emerging transcriptomic resources to study other iridoid and/or MIA producing plants can be further explored. The databases generated from large-scale

transcriptomic projects are excellent sources for searching candidate genes that may be involved in MIA biosynthesis. Studies of other MIA producing plants such as Apocynaceae family would be valuable in determining if the spatial separation and cellular localization of the MIA pathway is unique in *C. roseus* or common among the members of this family as it will give us more information of the architecture and the complexities involved in plant secondary metabolism.

References

Aerts, R. and De Luca, V. (1992). Phytochrome is involved in light-regulation of vindoline biosynthesis in *Catharanthus*. *Plant Physiology* **100**, 1029-1032.

Aerts, R., Schafer, A., Hesse, M., Baumann, T., and Slusarenko, A. (1996). Signaling molecules and the synthesis of alkaloids in *Catharanthus roseus* seedlings. *Phytochemistry* **42**, 417-422.

Ayora-Talavera, T., Chappell, J., Lozoya-Gloria, E., and Loyola-Vargas, V.M. (2002). Overexpression in *Catharanthus roseus* hairy roots of a truncated hamster 3-hydroxy-3-methylglutaryl-CoA reductase gene. *Applied Biochemistry and Biotechnology* **97**, 135-145.

Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., and Stöckigt, J. (2007). Molecular architecture of strictosidine glucosidase: The gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *The Plant Cell* **19**, 2886-2897.

Bergomaz, B. and Boppré, M. (1986). A simple insect diet for rearing Arctiidae and other moths. *Journal of Lepidopterist Society* **40**, 131–137.

Bernhardt, P., McCoy, E., and O'Connor, S. E. (2007). Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle. *Chemistry and Biology* **14**, 888-897.

Blom, T.J.M., Sierra, M., van Vliet, T.B., Franke-van Dijk, M.E.I., de Koning, P., van Iren, F., Verpoorte, R., and Libbenga, K. R. (1991). Uptake and accumulation of ajmalicine into isolated vacuoles of cultured cells of

Catharanthus roseus (L.) G. Don. and its conversion into serpentine. *Planta* **183**, 170-177.

Bowman, R.N. (1996). US Patent 5491285.

Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.

Burch-Smith, T.M., Anderson, J.C., Martin, G.B., and Dinesh-Kumar, S.P. (2004). Applications and advantages of virus-induced gene silencing for gene function studies in plants. *The Plant Journal* **39**, 734-746.

Burlat, V., Oudin, A., Courtois, M., Rideau, M., and St-Pierre, B. (2004). Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-drived primary metabolites. *The Plant Journal* **38**, 131-141.

Canel, C., Lopes-Cardoso, M.I., Whitmer, S., van der Fits, L., Pasquali, G., van der Heijden, R., Hoge, J. H., and Verpoorte, R. (1998). Effects of overexpression of strictosidine synthase and tryptophan decarboxylase on alkaloid production by cell cultures of *Catharanthus roseus*. *Planta* **205**, 414-419.

Chahed, K., Oudin, A., Guivarc'h, N., Hamdi, S., Chenieux, J., Rideau, M., and Clastre, M. (2000). 1-Deoxy-D-xylulose 5-phosphate from periwinkle: cDNA identification and induced gene expression in terpenoid indole alkaloid-producing cells. *Plant Physiology and Biochemistry* **38**, 559-566.

Chatel, G., Montiel, G., Pré, M., Memelink, J., Thiersault, M., St-Pierre, B., Doireau, P., and Gantet, P. (2003). CrMYC1, a *Catharanthus roseus* elicitor- and jasmonate-responsive bHLH transcription factor that binds the G-box element of the strictosidine synthase gene promoter. *Journal of Experimental Botany* **54**, 2587-2588.

Chen, S., Galan, M. C., Coltharp, C., and O'Connor, S.E. (2006). Redesign of a central enzyme in alkaloid biosynthesis. *Chemistry and Biology* **13**, 1137-1141.

Collu, G., Unver, N. , Peltenburg-Looman, A. M., van der Heijden, R., Verpoorte, R. , and Memelink, J. (2001). Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Letters* **508**, 215-20.

Contin, A., van der Heijden, R., Lefeber, A.W., and Verpoorte, R. (1998). The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture. *FEBS Letters* **434**, 413–416.

Costa, M.M., Hilliou, F., Duarte, P., Pereira, L.G., Almeida, L., Leech, M., Memelink, J., Barceló, A. R., and Sottomayor, M. (2008). Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiology* **146**, 403-417.

Courdavault, V., Thiersault, M., Courtois, M., Gantet, P., Oudin, A., Doireau, P., St-Pierre, B., and Giglioli-Guivarc'h, N. (2005). CaaX-prenyl transferases are essential for expression of genes involved in the early stages of monoterpenoid biosynthetic pathway in *Catharanthus roseus* cells. *Plant Molecular Biology* **57**, 855-870.

De Carolis, E., Chan, F., Balsevich, J., and De Luca, V. (1990). Isolation and characterization of a 2-oxoglutarate-dependent dioxygenase involved in the second-to-last step in vindoline biosynthesis. *Plant Physiology* **94**, 1323-1329.

De Carolis, E., and De Luca, V. (1993). Purification, characterization, and kinetic analysis of a 2-oxoglutarate-dependent dioxygenase involved in vindoline biosynthesis from *Catharanthus roseus*. *The Journal of Biological Chemistry* **268**, 5504-5511.

De Luca, V., Balsevich, J., Tyler, R.T., Eilert, U., Panchuk, B.D., and Kurz, W.G.W. (1986). Biosynthesis of indole alkaloids: developmental regulation of the biosynthetic pathway from tabersonine to vindoline in *Catharanthus roseus*. *Journal of Plant Physiology* **125**, 147-56.

De Luca, V. and Cutler, A.J. (1987). Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiology* **85**, 1099-1102.

De Luca, V., Marineau, C., and Brisson, N. (1989). Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2582-2586.

De Luca, V., Salim, V., Atsumi, S. M., and Yu, F. (2012a). Mining the biodiversity of plants: A revolution in the making. *Science* **336**, 1658-1661.

De Luca, V., Salim, V., Levac, D., Atsumi, S. M., and Yu, F. (2012b). Discovery and functional analysis of monoterpenoid indole alkaloid pathways in plants. *Methods in Enzymology* **515**, 207-229.

De Luca, V. and St-Pierre, B. (2000). The cell and developmental biology of alkaloid biosynthesis. *Trends Plant Science* **4**, 168-173.

Desgagné-Penix, I. and Facchini, P. J. (2012). Systematic silencing of benzylisoquinoline alkaloid biosynthetic genes reveals the major route to papaverine in opium poppy. *The Plant Journal* **72**, 331-344.

Deshpande, S.G., Joseph, M., and Sharma, R.N. (1988). Insect growth and development inhibition properties of *Catharanthus roseus*. *International Journal of Tropical Agricultural* **6**, 287–290.

Dethier, M. and De Luca, V. (1993). Partial purification of an *N*-methyltransferase involved in vindoline biosynthesis in *Catharanthus roseus*. *Phytochemistry* **32**, 673-678.

deWaal, A., Meijer, A., and Verpoorte, R. (1995). Strictosidine synthase from *Catharanthus roseus*: purification and characterization of multiple forms. *Biochemistry Journal* **306**, 571-580.

Dinesh-Kumar, S.P., Anandalakshmi, R., Marathe, R., Schiff, M., and Liu, Y. (2003). Virus-induced gene silencing. *Methods in Molecular Biology* **236**, 287–294.

Di Stilio, V.S., Kumar, R.A., Oddone, A.M., Tolkin, T.R., Salles, P., and McCarty, K. (2010) Virus-induced gene silencing as a tool for comparative functional studies in *Thalictrum*. *PLoS One* **5**: e12064.

Dinda, B., Debnath, S., and Banik, R. (2011). Naturally occurring iridoids and secoiridoids. An updated review, part 4. *Chem. Pharm. Bull* **59**, 803-833

Dinda, B., Debnath, S., and Harigaya, Y. (2007). Naturally occurring iridoids. A Review, part 1. *Chem. Pharm. Bull* **55**, 159-222.

Eisenreich, W., Rohdich, F., and Bacher, A. (2001). Deoxyxylulose phosphate pathway to terpenoids. *Trends in Plant Science* **6**, 78-84.

El-Sayed, M., Choi, Y. H., Frédérich, M., Roytrakul, S., and Verpoorte, S. (2004). Alkaloid accumulation in *Catharanthus roseus* cell suspension cultures fed with stemmadenine. *Biotechnology Letters* **26**, 793-798.

El-Sayed, M. and Verpoorte, R. (2007). *Catharanthus* terpenoid indole alkaloids: biosynthesis and regulation. *Phytochemistry Reviews* **6**, 277-305.

Facchini, P. J. (2001). Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering. *Annual Reviews of Plant Physiology and Plant Molecular Biology* **52**, 29-66.

Facchini, P. J., Bohlmann, J., Covello, P. S., De Luca, V., Mahadevan, R., Page, J. E., Ro, D.K., Sensen, C. W., Storms, R., and Martin, V.J. (2012). Synthetic biosystems for the production of high-value plant metabolites. *Trends in Biotechnology* **30**, 127-131.

Facchini, P.J. and De Luca, V. (2008). Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. *The Plant Journal* **54**, 763-784.

Fahn, W., Laussermair, E., Deus-Neumann, B., and Stöckigt, J. (1985). Late enzymes of vindoline biosynthesis *S*-adenosyl-L-methionine: 11-*O*-demethyl-17-*O*-deacetylvindoline 11-*O*-methylase and unspecific acetylesterase. *Plant Cell Reports* **4**, 337-340.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., and Gierl,

A. (1997). Analysis of a chemical plant defense mechanism in grasses. *Science* **277**, 696-699.

Fukushima, E.O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., and Muranaka, T. (2013). Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant Cell Physiology*: doi: 10.1093/pcp/pct015.

Gamborg, O.L., Murashige, T., Thorpe, T.A., and Vasil, I.K. (1976). Plant tissue culture media. *In Vitro* **12**, 473–478.

Geerlings, A., Ibanez, M., Memelink, J., van der Heijden, R., and Verpoorte, R. (2000). Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *The Journal of Biological Chemistry* **275**, 3051-3056.

Gerasimenko, I., Sheludko, Y., Ma, X., and Stöckigt, J. (2002). Heterologous expression of a *Rauvolfia* cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids. *European Journal of Biochemistry* **269**, 2204-2213.

Geu-Flores, F., Sherden, N.H., Courdavault, V., Burlat, V., Glenn, W.S., Wu, C., Nims, E., Cui, Y., and O'Connor, S.E. (2012). An alternative route to cylic terpenes by reductive cyclization in iridoid biosynthesis. *Nature* **492**, 138-142.

Giddings, L.A., Liscombe, D.K., Hamilton, J.P., Childs, K.L., DellaPenna, D., Buell, C.R., and O'Connor, S.E. (2011). A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in *Catharanthus roseus*. *The Journal of Biological Chemistry* **286**, 16751-16757.

Gietz, D., St. Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation in intact yeast cells. *Nucleic Acids Research* **20**, 1425.

Gniwotta, F., Vogg, G., Gartmann, V., Carver, T.L.W., Riederer, W., and Jetter, R. (2005). What do microbes encounter at the plant surface? Chemical composition of pea leaf cuticular waxes. *Plant Physiology* **139**, 519–530.

Góngora-Castillo, E., Childs, K. L., Fedewa, G., Hamilton, J. P., Liscombe, D. K., Magallanes-Lundback, M., Mandadj, K.K., Nims, E., Runguphan, W., Vaillancourt, B., Varbanova-Herde, M., DellaPenna, D., McKnight, T. D., O'Connor, S., and Buell, C. R. (2012). Development of transcriptomic resources for interrogating the biosynthesis of monoterpene indole alkaloids in medicinal plant species. *PLoS ONE*, **7** (12), e52506.

Gould, B. and Kramer, E.M. (2007). Virus-induced gene silencing as a tool for functional analyses in the emerging model plant *Aquilegia* (columbine, Ranunculaceae). *Plant Methods* **3**, 6.

Gueritte, F., Bac, N. V., Langlois, Y., and Potier, P. (1980). Biosynthesis of antitumour alkaloids from *Catharanthus roseus*. Conversion of 20'-deoxyleurosidine into vinblastine. *Journal of the Chemical Society, Chemical Communications*, 452-453.

Guirimand, G., Burlat, V., Oudin, A., Lanoue, A., St-Pierre, B., and Courdavault, V. (2009). Optimization of the transient transformation of *Catharanthus roseus* cells by particle bombardment and its application to the subcellular localization of hydroxymethylbutenyl 4-diphosphate synthase and geraniol 10-hydroxylase. *Plant Cell Reports* **28**, 1215-1234.

Guirimand G., Courdavault V., Lanoue A., Mahroug S., Guihur A., Blanc N., Giglioli Guivarc'h N., St-Pierre B., and Burlat V. (2010). Strictosidine activation in Apocynaceae: towards a "nuclear time bomb"? *BMC Plant Biology* **10**, 182.

Guirimand, G., Guihur, A., Ginis, O., Poutrain, P., Héricourt, F., Oudin, A., Lanoue, A., St-Pierre, B., Burlat, V., and Courdavault, V. (2011a). The subcellular organization of strictosidine biosynthesis in *Catharanthus roseus* epidermis highlights several trans-tonoplast translocations of intermediate metabolites. *FEBS Journal* **278**, 749-763.

Guirimand, G., Guihur, A., Poutrain, P., Hericourt, F., Mahroug, S., St-Pierre, B., Burlat, V., and Courdavault, V. (2011b). Spatial organization of the vindoline biosynthetic pathway in *Catharanthus roseus*. *Journal of Plant Physiology* **168**, 549-557.

Hagel, J.M. and Facchini, P.J. (2010). Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nature Chemical Biology* **6**, 273-755.

Hasada, K., Yoshida, T., Yamazaki, T., Sugimoto, N., Nishimura, T., Nagatsu, A., and Mizukami, H. (2011). Application of ^{1}H-NMR spectroscopy to validation of berberine alkaloid reagents and to chemical evaluation of *Coptis Rhizoma*. *Journal of Natural Medicines* **65**, 262-267.

Hemscheidt, T. and Zenk, M. H. (1985). Partial purification and characterization of a NADPH dependent tetrahydroalstonine synthase from *Catharanthus roseus* cell cultures. *Plant Cell Reports* **4**, 216-219.

Hileman, L.C., Drea, S., Martino, G., Litt, A., and Irish, V.F. (2005). Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (opium poppy). *The Plant Journal* **44**, 334-341.

Hong, S.B., Peebles, C.A., Shanks, J.V., San, K.Y., and Gibson, S.I. (2006). Expression of the *Arabidopsis* feedback-insensitive anthranilate synthase holoenzyme and tryptophan decarboxylase genes in *Catharanthus roseus* hairy roots. *Journal of Biotechnology* **122**, 28-38.

Hotze, M., Schröder, G., and Schröder, J. (1995). Cinnamate 4-hydroxylase from *Catharanthus roseus* and a strategy for the functional expression of plant cytochrome P_{450} proteins as translational fusions with P_{450} reductase in *Escherichia coli*. *FEBS Letters* **374**, 345-350.

Hughes, E.H., Hong, S.B., Gibson, S.I., Shanks, J.V., and San, K.Y. (2004). Metabolic enginerring of the indole pathway in *Catharanthus roseus* hairy roots and increased accumulation of tryptamine and serpentine. *Metabolic Engineering* **6**, 268-276.

Ikeda, H., Esaki, N., Nakai, S., Hashimoto, K., Uesato, S., Soda, K., and Fujita, T. (1991). Acyclic monoterpene primary alcohol: $NADP^+$ oxidoreductase of *Rauwolfia serpentina* cells: The key enzyme in biosynthesis of monoterpene alcohols. *Journal of Biochemistry* **109**, 341-347.

Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J., Strack, D., Matern, U., and Schröder, J. (2000). Indole alkaloid biosynthesis in *Catharanthus roseus*: New enzyme activities and identification of cytochrome P450 *CYP72A1* as secologanin synthase. *The Plant Journal* **24**, 797-804.

Ishikawa, H., Colby, D.A., and Boger, D. L. (2008). Direct coupling of catharanthine and vindoline to provide vinblastine: Total synthesis of (+)- and ent-(-)-vinblastine. *Journal of the American Chemical Society* **130**, 420-421.

Kaltenbach, M., Schröder, G., Schmelzer, E., Lutz, V., and Schröder, J. (1999). Flavonoid hydroxylase from *Catharanthus roseus*: cDNA, heterologous expression, enzyme properties and cell-type specific expression in plants. *The Plant Journal* **19**, 183-193.

Kaminaga, Y., Sahin, F.P., and Mizukami, H. (2004). Molecular cloning and characterization of a glucosyltransferase catalyzing glucosylation of curcumin in cultured *Catharanthus roseus* cells. *FEBS Letters* **567**, 197-202.

Katano, N., Yamamoto, H., Iio, R., and Inoue, K. (2001). 7-deoxyloganin 7-hydroxylase in *Lonicera japonica* cell cultures. *Phytochemistry* **58**, 53-58.

Kawai, H., Kuroyanagi, M., and Ueno, A. (1988). Iridoid glucosides from *Lonicera japonica* THUNB. *Chemical and Pharmaceutical Bulletin* **36**, 3664-3666.

Kubo, I. and Matsumoto, A. (1984). Secreted oleanolic acid on the cuticle *Olea europaea* (Oleaceae); a chemical barrier to fungal attack. *Experientia* **40**, 937–938.

Kuboyama, T., Yokoshima, S., Tokuyama, H., and Fukuyama, T. (2004). Stereocontrolled total synthesis of (+)-vincristine. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11966-11970.

Kuehne, M. E., Matson, P. A., and Bornmann, W. G. (1991). Enantioselective syntheses of vinblastine, leurosidine, vincovaline and 20'-epi-vincovaline. *Journal of Organic Chemistry* **56**, 513-528.

Kunert, M., Søe, A., Bartram, S., Discher, S., Tolzin-Banasch, K., Nie, L., David, A., Pasteels, J., and Boland, W. (2008). De novo biosynthesis versus sequestration: A network of transport systems supports in iridoid producing leaf beetle larvae both modes of defense. *Insect Biochemical Molecular Biology* **38**, 895–904.

Kutchan, T.M., Hampp, N., Lottspeich, F., Beyreuther, K., and Zenk, M.H. (1988). The cDNA clone for strictosidine synthase from *Rauwolfia serpentina*: DNA sequence determination and expression in *Escherichia coli*. *FEBS Letters* **237**, 40-44.

Kutney J. P., Choi, L. S. L., Nakano, J., Tsukamoto, H., McHugh, M., and Boulet C. A. (1988). A highly efficient and commercially important synthesis of the antitumor catharanthus alkaloids vinblastine and leurosidine from catharanthine and vindoline. *Heterocycles* **27**, 1845-1853.

Laflamme, P., St-Pierre, B., and De Luca, V. (2001). Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-O-acetyltransferase. *Plant Physiology* **125**, 189–198.

Lange, B. M. and Croteau, R. (1999). Isopentenyl diphosphate biosynthesis via a mevalonate-independent pathway : isopentenyl monophosphate kinase catalyzes the terminal enzymatic step. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13714-13719.

Lange, B. M., Wildung, M, R., Stauber, E. J., Sanchez, C., Pouchnik, D., and Croteau, R. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes.

Proceedings of the National Academy of Sciences of the United States of America **97**, 2934-2939.

Lee, E.J. and Facchini, P. (2010). Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. *The Plant Cell* **22**, 3489-3503.

Levac, D., Murata, J., Kim, W.S., and De Luca, V. (2008). Application of carborundum abrasion for investigating the leaf epidermis: Molecular cloning of *Catharanthus roseus* 16- hydroxytabersonine-16-O-methyltransferase. *The Plant Journal* **53**, 225–236.

Li, R., Reed, D.W., Liu, E., Nowak, J., Pelcher, L.E., Page, J.E., and Covello, P.S. (2006). Functional genomic analysis of alkaloid biosynthesis in *Hyoscyamus niger* reveals a cytochrome P450 involved in littorine rearrangement. *Chemistry and Biology* **13**, 513-520.

Linsmaier, E.M. and Skoog, F. (1965). Organic growth factor requirement of tobacco tissue cultures. *Physiologia Plantarum* **18**, 100-127.

Liscombe, D. and O'Connor, S. (2011). A virus-induced gene silencing approach to understanding alkaloid metabolism in *Catharanthus roseus*. *Phytochemistry* **72**, 1969-1977.

Liscombe, D.K., Usera, A.R., and O'Connor, S.E. (2010). Homolog of tocopherol C-methyltransferases catalyzes *N*-methylation in anticancer alkaloid biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18793-18798.

Liu, Y., Schiff, M., Marathe, R., and Dinesh-Kumar, S.P. (2002). Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for *N*-mediated resistance to tobacco mosaic virus. *The Plant Journal* **30**, 415-429.

Lopez-Meyer, M. and Nessler, C.L. (1997). Tryptophan decarboxylase is encoded by two autonomously regulated genes in *Camptotheca acuminata* which are differentially expressed during development and stress. *The Plant Journal* **11**, 1167-1175.

Luijendijk, T.J.C., Stevens, L.H., and Verpoorte, R. (1998). Purification and characterization of strictosidine beta-D-glucosidase from *Catharanthus roseus* cell suspension cultures. *Plant Physiology and Biochemistry* **36**, 419-425.

Luijendijk, T.J.C., Van der Meijden, E., and Verpoorte, R. (1996). Involvement of strictosidine as a defensive chemical in *Catharanthus roseus*. *Journal of Chemical Ecology* **22**, 1355-1366.

Ma, X., Panjikar, S., Koepke, J., Loris, E., and Stöckigt, J. (2006). The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. *The Plant Cell* **18**, 907-920.

Madyastha, K.M., Guarnaccia, R., Baxter, C., and Coscia, C.J. (1973). *S-Adenosyl-L-methionine*: loganic acid methyltransferase. *The Journal of Biological Chemistry* **248**, 2497-2501.

Madyastha, K.M., Ridgway, J.E., Dwyer, J.G., and Coscia, C.J. (1977). Subcellular localization of a cytochrome P-450-dependent monooxygenase in vesicles of the higher plant *Catharanthus roseus*. *Journal of Cell Biology* **72**, 303-313.

Magnotta, M., Murata, J., Chen, J., and De Luca, V. (2006). Identification of a low vindoline accumulating cultivar of *Catharanthus roseus* (L.) G. Don by alkaloid and enzymatic profiling. *Phytochemistry* **67**, 1758-1764.

Magnotta, M., Murata, J., Chen, J., and De Luca, V. (2007). Expression of deacetylvindoline-4-O-acetyltransferase in *Catharanthus roseus* hairy roots. *Phytochemistry* **68**, 1922-1931.

Magnus, P., Mendoza, J.S., Stamford, A., Ladlow, M., and Willis, P. (1992). Nonoxidative coupling methodology for the synthesis of the antitumor bisindole alkaloid vinblastine and a lower-half analog: Solvent effect on the stereochemistry of the crucial C-15/C-18'bond. *Journal of American Chemical Society* **114**, 10232–10245.

Mahroug, S., Burlat, V., and St-Pierre, B. (2007). Cellular and sub-cellular organisation of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*. *Phytochemistry Reviews* **6**, 363–381.

Mangeney, P., Andriamialisoa, R.Z., Langlois, N., Langlois, Y., and Potier, P. (1979). Preparation of vinblastine, vincristine, and leurosidine, antitumor alkaloids from *Catharanthus* species (Apocynaceae). *Journal of American Chemical* Society **101**, 2243 -2245.

Maresh, J.J., Giddings, L.A., Friedrich, A., Loris, E.A., Panjikar, S., Trout, B.L., Stöckigt, J., Peters, B., and O'Connor, S.E. (2008). Strictosidine synthase: Mechanism of a Pictet-Spengler catalyzing enzyme. *Journal of the American Chemical Society* **130**, 710-723.

Masada, S., Terasaka, K., Oguchi, Y., Okazaki, S., Mizushima, T., and Mizukami, H. (2009). Functional and structural characterization of a flavonoid glucoside 1,6-glucosyltransferase from *Catharanthus roseus*. *Plant and Cell Physiology* **50**, 1401-1415.

McCoy, E. and O'Connor, S. E. (2006). Directed biosynthesis of alkaloid analogs in the medicinal plant *Catharanthus roseus*. *Journal of the American Chemical Society* **128**, 14276-14277.

McKnight, T.D., Bergey, D.R., Burnett, R.J., and Nessler, C.L. (1991). Expression of enzymatically active and correctly targeted strictosidine synthase in transgenic tobacco plants. *Planta* **185**, 148-52.

McKnight, T.D., Roessner, C.A., Devagupta, R., Scott, A.I., and Nessler, C.L. (1990). Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*. *Nucleic Acids Research* **18**, 4939.

Meehan, T. D. and Coscia, C. J. (1973). Hydroxylation of geraniol and nerol by a monooxygenase from *Vinca rosea*. *Biochemical and Biophysical Research Communications* **53**, 1043-1048.

Meisner, J., Weissenberg, M., Palevitch, D., and Aharonson, N. (1981). Phagodeterrency induced by leaves and leaf extracts of *Catharanthus roseus* in the larvae of *Spodoptera littoralis*. *Journal of Economic Entomology* **74**, 131–135.

Memelink, J. and Gantet, P. (2007). Transcription factors involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Phytochemistry Reviews* **6**, 353-362.

Menke, F., Parchmann, S., Mueller, M., Kijne, J., and Memelink, J. (1999). Involvement of the octadecanoid pathway and protein phosphorylation in fungal

elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus*. *Plant Physiology* **119**, 1289-1296.

Mizutani, M. and Sato, F. (2011). Unusual P450 reactions in plant secondary metabolism. *Archives of Biochemistry and Biophysics* **507**, 194-203.

Mizutani, M. and Ohta, D. (2010). Diversification of P450 genes during land plant evolution. *Annual Reviews of Plant Biology* **61**, 291-315.

Murata J., Bienzle, D., Brandle, J.E., Sensen, C.W., and De Luca, V. (2006). Expressed sequence tags from Madagascar periwinkle (*Catharanthus roseus*). *FEBS Letters* **580**, 4501-4507.

Murata, J. and De Luca, V. (2005). Localization of tabersonine 16-hydroxylase and 16-OH-tabersonine-16-*O*-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *The Plant Journal* **44**, 581-594.

Murata, J., Roepke, J., Gordon, H., and De Luca, V. (2008). The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. *The Plant Cell* **20**, 524-542.

Naaranlahti, T., Auriola, S., and Lapinjoki, S.P. (1991). Growth-related dimerization of vindoline and catharanthine in *Catharanthus roseus* and effect of wounding on the process. *Phytochemistry* **30**, 1451–1453.

Nagatoshi, M., Terasaka, K., Nagatsu, A., and Mizukami, H. (2011). Iridoid-specific glucosyltransferase from *Gardenia jasminoides*. *Journal of Biological Chemistry*. **286**, 32866-32874.

Nelson, D. R., Ming, R., Alam, M., and Schuler, M. A. (2008). Comparison of cytochrome P450 genes from six plant genomes. *Tropical Plant Biology* **1**, 216-235.

Newman, J. D. and Chappell, J. (1999). Isoprenoid biosynthesis in plants: carbon partitioning within the cytoplasmic pathway. *Critical Reviews in Biochemistry and Molecular Biology* **34**, 95-106.

O'Connor, S. E. (2012). Strategies for engineering plant natural products: The iridoid-derived monoterpene indole alkaloids of *Catharanthus roseus*. *Methods in Enzymology* **515**, 189-206.

O'Connor, S. E. and Maresh, J. J. (2006). Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Natural Product Reports* **23**, 532-547.

Oudin, A., Courtois, M., Rideau, M., and Clastre, M. (2007a). The iridoid pathway in *Catharanthus roseus* alkaloid biosynthesis. *Phytochemistry Reviews* **6**, 259-276.

Oudin, A., Mahroug, S., Courdavault, V., Hervouet, N., Zelwer, C., Rodríguez-Concepcíon, M., St-Pierre, B., and Burlat, V. (2007b). Spatial distribution and hormonal regulation of gene products from methyl erythritol phosphate and monoterpene-secoiridoid pathways in *Catharanthus roseus*. *Plant Molecular Biol*ogy **65**, 13-30.

Papon, N., Vansiri, A., Gantet, P., Chenieux, J.C., Rideau, M., and Creche, J. (2004). Histidine-containing phosphotransfer domain extinction by RNA interference turns off a cytokinin signaling circuitry in *Catharanthus roseus* suspension cells. *FEBS Letters* **558**, 85-88.

Pasquali, G., Gooddijn, O.J.M., de Waal, A., Verpoorte, A., Schilperoort, R.A., Hoge, J.H., and Memelink, J. (1992). Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant Molecular Biology* **18**, 1121-1131.

Pasquali, G., Porto, D.D., and Fett-Neto, A.G. (2006). Metabolic engineering of cell cultures versus whole plant complexity in production of bioactive monoterpene indole alkaloids: recent progress related to an old dilemma. *Journal of Bioscience and Bioengineering* **101**, 287-296.

Pauw, B., Hilliou, F.A.O., Sandonis, V., Chatel, M.G., de Wolf, C.J.F., Champion, A., Pré, M., van Duijn, B., Kijne, J.W., van der Fits, L., and Memelink, J. (2004). Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. *The Journal of Biological Chemistry* **279**, 52940-52948.

Peñuelas, J., Sardans, J., Stefanescu, C., Parella, T., and Filella, I. (2006). *Lonicera implexa* leaves bearing naturally laid eggs of the specialist herbivore *Euphydryas aurinia* have dramatically greater concentrations of iridoid glycosides than other leaves. *Journal of Chemical Ecology* **32**, 1925-1933.

Pérez-Bonilla, M., Salido, S., van Beek, T.A., de Waard, P., Linarez-Palomino, P.J., Sánchez, A., and Altarejos, J. (2011). Isolation of antioxidative secoiridoids from olive wood (*Olea europaea* L.) guided by on-line HPLC-DAD-radical scavenging detection. *Food Chemistry* **124**, 36-41.

Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996). Yeast expression of animal and plant P450s in optimized redox environments. *Methods in Enzymology* **272**, 51-64.

Power, R., Kurz, W.G.W., and De Luca, V. (1990). Purification and characterization of acetylcoenzyme A: deacetylvindoline 4-*O*-acetyltransferase from *Catharanthus roseus*. *Archives in Biochemistry and Biophysics* **279**, 370-376.

Prakash, V. and Timasheff, S.N. (1991). Mechanism of interaction of vinca alkaloids with tubulin: Catharanthine and vindoline. *Biochemistry* **30**, 873–880.

Qureshi, A.A. and Scott, A.I. (1968). Interconversion of corynanthe, aspidosperma, and iboga alkaloids a model for indole alkaloid biosynthesis. *Chemical Communications* 945-946.

Ramos-Valvidia, A.C., van der Heijden, R., and Verpoorte, R. (1997). Isopentenyl diphosphate isomerase: a core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. *Natural Product Reports* **14**, 591-603.

Ratcliff, F., Martin-Hernandez, A.M., and Baulcombe, D.C. (2001). Tobacco rattle virus as a vector for analysis of gene function by silencing. *The Plant Journal* **25**, 237-245.

Rischer, H., Oresic, M., Seppanen-Laakso, T., Katajamaa, M., Lammertyn, F., Ardilles-Diaz, W., van Montagu, M.C., Inzé, D., Oksman-Caldentey, K.M., and Goossens, A. (2006). Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5614-19.

Ro, D.K., Ehlting, J., and Douglas C.J. (2002). Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar. *Plant Physiology* **130**, 1837-1851.

Ro, D.K., Ouellet, M., Paradise, E.M., Burd, H., Eng, D., Paddon, C.J., Newman, J.D., and Keasling, J.D. (2008). Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Technology* **8**: 83

Ro, D.K., Paradise, E. M., Ouellet, M., Fisher, K. J., and Newman, K. L. (2006). Production of the antimalarial drug precursor artermisinic acid in engineered yeast. *Nature* **440**, 940-943.

Rodriguez, S., Compagnon, V., Crouch, N. P., St-Pierre, B., and De Luca, V. (2003). Jasmonate-induced epoxidation of tabersonine by cytochrome P-450 in hairy root cultures of *Catharanthus roseus*. *Phytochemistry* **64**, 401-409.

Rodriguez-Conception, A. and Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiology* **130**, 1079-1089.

Roepke, J., Salim, V., Wu, M., Thamm, A.M., Murata, J., Ploss, K., Boland, W., and De Luca V. (2010). Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15287-15292.

Roewer, I., Cloutier, N., Nessler, C., and De Luca, V. (1992). Transient induction of tryptophan decarboxylase (TDC) and strictosidine synthase (SS) genes in cell suspension cultures of *Catharanthus roseus*. *Plant Cell Reports* **11**, 86-89.

Rohdich, F., Kis, K., Bacher, A., and Eisenreich, W. (2001). The non-mevalonate pathway of isoprenoids: genes, enzymes and intermediates. *Current Opinion in Chemical Biology* **5**, 535-540.

Rohmer, M. (1999). The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae, and higher plants. *Natural Product Reports* **16**, 565-574.

Rotenberg, D., Thompson, T.S., German, T. L., and Willis, D. K. (2006). Methods for effective real-time RT-PCR analysis of virus-induced gene silencing. *J. Virological Methods* **138**, 49-59.

Runguphan, W., Maresh, J.J., and O'Connor, S.E. (2009). Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13673-13678.

Runguphan, W., Qu, X., and O'Connor, S. E. (2010). Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **468,** 461-464.

Salim, V. and De Luca, V. (2013). Towards complete elucidation of monoterpene indole alkaloid biosynthesis pathway: *Catharanthus roseus* as a pioneer system. *Advances in Botanical Research* **68**, 1-37.

Samuels, L., Kunst, L., and Jetter, R. (2008). Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annual Reviews of Plant Biology* **59**, 683–707.

Sanchez-Iturbe, P., Galaz-Avalos, R., and Loyola-Vargas, V. (2005). Determination of partial purification of 10-oxogeranial: iridodial cyclase an enzyme catalyzing the synthesis of iridodial from *Catharanthus roseus* hairy roots. *Phyton* **54,** 55-69.

Schilmiller, A. L., Last, R. L., and Pichersky, E. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. *The Plant Journal* **54**, 702-711.

Schröder, G., Unterbusch, E., Kaltenbach, M., Schmidt, J., Strack, D., De Luca V., and Schröder, J. (1999). Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase. *FEBS Letters* **458**, 97-102.

Seki, H., Sawai, S., Ohyama, K., Mizutani, M., Ohnishi, T., Sudo, H., Fukushima, E.O., Akashi, T., Aoki, T., Saito, K., and Muranaka, T. (2011). Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *The Plant Cell* **23,** 4112-4123.

Senthil-Kumar, M. and Mysore, K.S. (2011). New dimension for VIGS in plant functional genomics. *Trends in Plant Science* **16**, 656-665.

Shukla, A.K., Shasany, A.K., Gupta, M.M., and Khanuja, S.P. (2006). Transcriptome analysis in *Catharanthus roseus* leaves and roots for comparative terpenoid indole alkaloid profiles. *Journal of Experimental Botany* **57**, 3921-3932.

Siberil, Y., Benhamron, S., Memelink, J., Giglioli-Guivarc'h, N., Thiersault, M., Boisson, B., Doireau, P., and Gantet, P. (2001). *Catharanthus roseus* G-box

binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. *Plant Molecular Biology* **45**, 477-488.

Simkin, A. J., Miettinen, K., Claudel, P., Burlat, V., Guirimand, G., Courdavault, V., Papon, N., Meyer, S., Godet, S., St-Pierre, B., Giglioli-Guivarc'h, N., Fischer, M. J. C., Memelink, J., and Clastre, M. (2013). Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. *Phytochemistry* **85**, 36-43.

Stevens, L.H., Blom, T.J.M., and Verpoorte, R. (1993). Subcellular localization of tryptophan decarboxylase, strictosidine synthase and strictosidine glucosidase in suspension-cultured cells of *Catharanthus roseus* and *Tabernaemontana divaricata*. *Plant Cell Reports* **12**, 573-576.

Stöckigt, J. and Zenk, M. (1977). Strictosidine (isovincoside): The key intermediate in the biosynthesis on monoterpene indole alkaloids. *Journal of Chemical Society, Chemical Communications*, 646-648.

St-Pierre, B. and De Luca, V. (1995). A cytochrome P-450 monooxygenase catalyzes the first step in the conversion of tabersonine to vindoline in *Catharanthus roseus*. *Plant Physiology* **109**, 131-139.

St-Pierre, B., Laflamme, P., Alarco, A.M., and De Luca, V. (1998). The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. *The Plant Journal* **14**, 703-713.

St-Pierre, B., Vázquez-Flota, F.A., and De Luca, V. (1999). Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. *The Plant Cell* **11**, 887-900.

Suttipanta, N., Pattanaik, S., Kulshrestha, M., Patra, B., Singh, S.K., and Yuan, L. (2011). The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiology* **157**, 2081-2093.

Szabo, L. (2008). Molecular evolutionary lines in the formation of indole alkaloids derived from secologanin. *Arkivoc* 167-181.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731-2739.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Treimer, J. and Zenk, M. (1979). Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation. *European Journal of Biochemistry* **101**, 225-233.

Uesato, S., Miyauchi, M., Itoh, H., and Inouye, H. (1986). Biosynthesis of iridoid glucosides in *Galium mollugo*, *G. spurium* var *Echinospermon* and *Deutzia*

crenata, intermediacy of deoxyloganic acid, loganin, and iridodial glucoside. *Phytochemistry* **25**, 2515-2521.

Uesato, S., Ikeda, H., Fujita, T., Inouye, H., and Zenk, M. (1987). Elucidation of iridodial formation mechanism: Partial purification and characterization of the novel monoterpene cyclase from *Rauwolfia serpentina* cell suspension cultures. *Tetrahedron Letters* **28**, 4431-4434.

Usia, T., Watabe, T., Kadota, S., and Tezuka, Y. (2005). Cytochrome P450 2D6 (CYP2D6) inhibitory constituents of *Catharanthus roseus*. *Biological and Pharmaceutical Bulletin* **28**, 1021-1024.

Valant-Vetschera, K,M. and Brem, B. (2006). Chemodiversity of exudate flavonoids as highlighted by publications of Eckhard Wollenweber. *Natural Product Communications* **1**, 921–926.

van der Fits, L. and Memelink, J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* **289**, 295-297.

van der Heijden, R., Jacobs, D. I., Snoeijer, W., Hallard, D., and Verpoorte, R. (2004). The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Current Medicinal Chemistry* **11**, 607-628.

Vazquez-Flota, F.A., De Carolis, E., Alarco, A.M., and De Luca, V. (1997). Molecular cloning and characterization of deacetoxyvindoline 4-hydroxylase, a 2-oxoglutarate dependent dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus* (L.) G. Don. *Plant Molecular Biology* **34**, 935-948.

Vazquez-Flota F. and De Luca, V. (1998a). Developmental and light regulation of desacetoxyvindoline 4-hydroxylase in *Catharanthus roseus* (L.) G. Don. *Plant Physiology* **117**, 1351-1361.

Vazquez-Flota, F. and De Luca, V. (1998b). Jasmonate modulates developmental and light regulated alkaloid biosynthesis in *Catharanthus roseus*. *Phytochemistry* **49**, 395-402.

Veau, B., Courtois, M., Oudin, A., Chenieux, J. C., Rideau, M., and Clastre, M. (2000). Cloning and expression of cDNAs encoding two enzymes of the MEP pathway in *Catharanthus roseus*. *Biochimica et Biophysica Acta* **1517**, 159-163.

Vogt, T. and Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in Plant Science* **5**, 380–386.

Vrieling, K. and Derridj, S. (2003). Pyrrolizidine alkaloids in and on the leaf surface of *Senecio jacobaea* L. *Phytochemistry* **64**, 1223–1228.

Wege, S., Scholz, A., Gleissberg, S., and Becker, A. (2007). Highly efficient virus-induced gene silencing (VIGS) in California poppy (*Eschscholzia californica*): an evaluation of VIGS as a strategy to obtain functional data from nonmodel plants. *Annals of Botany* **100**, 641-649.

Wijekoon, C.P. and Facchini, P.J. (2012). Systematic knockdown of morphine pathway enzymes in opium poppy using virus-induced gene silencing. *The Plant Journal* **69**, 1052-1063.

Xiao, M., Zhang, Y., Chen, X., Lee, E.J.., Barber, C., Chakrabarty, R., Desgagné-Penix, I., Haslam, T. M., Kim, Y.B., Liu, E., MacNevin, G., Masada-Atsumi, S., Reed, D., Stout, J. M., Zerbe, P., Zhang, Y., Bohlmann, J., Covello, P. S.,

De Luca, V., Page, J. E., Ro, D.K., Martin, V.J.J., Facchini, P. J., and Sensen, C. W. (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *Journal of Biotechnology* **166**, 122-134.

Yamamoto, H., Katano, N., Ooi, A., and Inoue, K. (1999). Transformation of loganin and 7-deoxyloganin into secologanin by *Lonicera japonica* cell suspension cultures. *Phytochemistry* **50**, 417-422.

Yamamoto, H., Katano, N., Ooi, A., and Inoue, K. (2000). Secologanin synthase which catalyzes the oxidative cleavage of loganin into secologanin is a cytochrome P450. *Phytochemistry* **53**, 7-12.

Yamazaki, Y., Sudo, H., Yamazaki, M., Aimi, N., and Saito, K. (2003). Camptothecin biosynthetic genes in hairy roots of *Ophiorrhiza pumila*: cloning, characterization and differential expression in tissues and by stress compounds. *Plant Cell and Physiology* **44**, 395-403.

Yokoshima, S., Ueda, T., Kobayashi, S., Sato, A., Kuboyama, T., Tokuyama, H., and Fukuyama, T. (2002). Stereocontrolled total synthesis of (+)-vinblastine. *Journal of American Chemical Society* **124**, 2137–2139.

Yordanov, M., Dimitrova, P., Patkar, S., Falcocchio, S., Xoxi, E., Saso, L., and Ivanovska, N. (2005). Ibogaine reduces organ colonization in murine systemic and gastrointestinal *Candida albicans* infections. *Journal of Medicinal Microbiology* **54**, 647–653.

Zárate, R. and Verpoorte, R. (2007). Strategies for the genetic modification of the medicinal plant *Catharanthus roseus* (L.) G. Don. *Phytochemistry Reviews* **6**, 475-491.

Zhang, H., Hedhili, S., Montiel, G., Zhang, Y., Chatel, G., Pré, M., Gantet, P., and Memelink, J. (2011). The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. *The Plant Journal* **67**, 61-71.

Zhao, J. and Verpoorte, R. (2007). Manipulating indole alkaloid production by *Catharanthus roseus* cell cultures in bioreactors: From biochemical processing to metabolic engineering. *Phytochemistry Reviews* **6**, 435-457.

Zhu, J., Guggisberg, A., Kalt-Hadamowsky, M., and Hesse, M. (1990). Chemotaxonomic study of the genus *Tabernaemontana* (Apocynaceae) based on their indole alkaloid content. *Plant Systematics and Evolution* **172**, 13-34.

Ziegler, J. and Facchini, P. J. (2008). Alkaloid biosynthesis: Metabolism and trafficking. *Annual Reviews of Plant Biology* **59**, 735-769.

Appendices

Appendix I

Chapter 2 Supplementary Tables:

Table 2.S1 Relative fresh *C. roseus* leaf consumption by various insect species during 1-wk feeding

Insect species	Relative leaf consumption	Days before death	Detection of MIAs in insect feces		
			Catharanthine	Vindoline	Anhydrovinblastine
S. littoralis third-instar larvae	++	ND	+	—	—
S. littoralis fourth- instar larvae	+++	ND	+	—	—
S. eridania third-instar larvae	—	ND	NS	NS	NS
S. eridania fourth-instar larvae	+	ND	NS	NS	NS
H. armigera third-instar larvae *a*	+	ND	NS	NS	NS
P. cochleariae third-instar larvae	—	5	NS	NS	NS
B. mori fifth-instar larvae	—	5	NS	NS	NS

All feeding studies were performed with *C. roseus* cv Roseus leaves except for those conducted with *Bombyx mori* where cv Little Delicata leaves were used.
ND, none detected; NS, not studied (these insects did not eat or they fed poorly on *Catharanthus* leaves and produced no feces).

Table 2.S2 Distribution of each alkaloids detected in various samples extracted from fifth-instar *B. mori* larvae fed with *C. roseus* leaves mixed Mulberry artificial diet in triplicate

Samples	Catharanthine, %	Vindoline, %	Anhydrovinblastine, %
Regurgitates	0.21 ± 0.02	0.31 ± 0.08	0
Feces	6.35 ± 1.38	6.14 ± 1.99	0
Gut	0.71 ± 0.06	1.18 ± 0.14	0
Body	4.07 ± 0.48	6.55 ± 1.51	0
Remaining prepared diet	75.19 ± 10.91	81.37 ± 7.92	0

Chapter 2 Supplementary Figures:

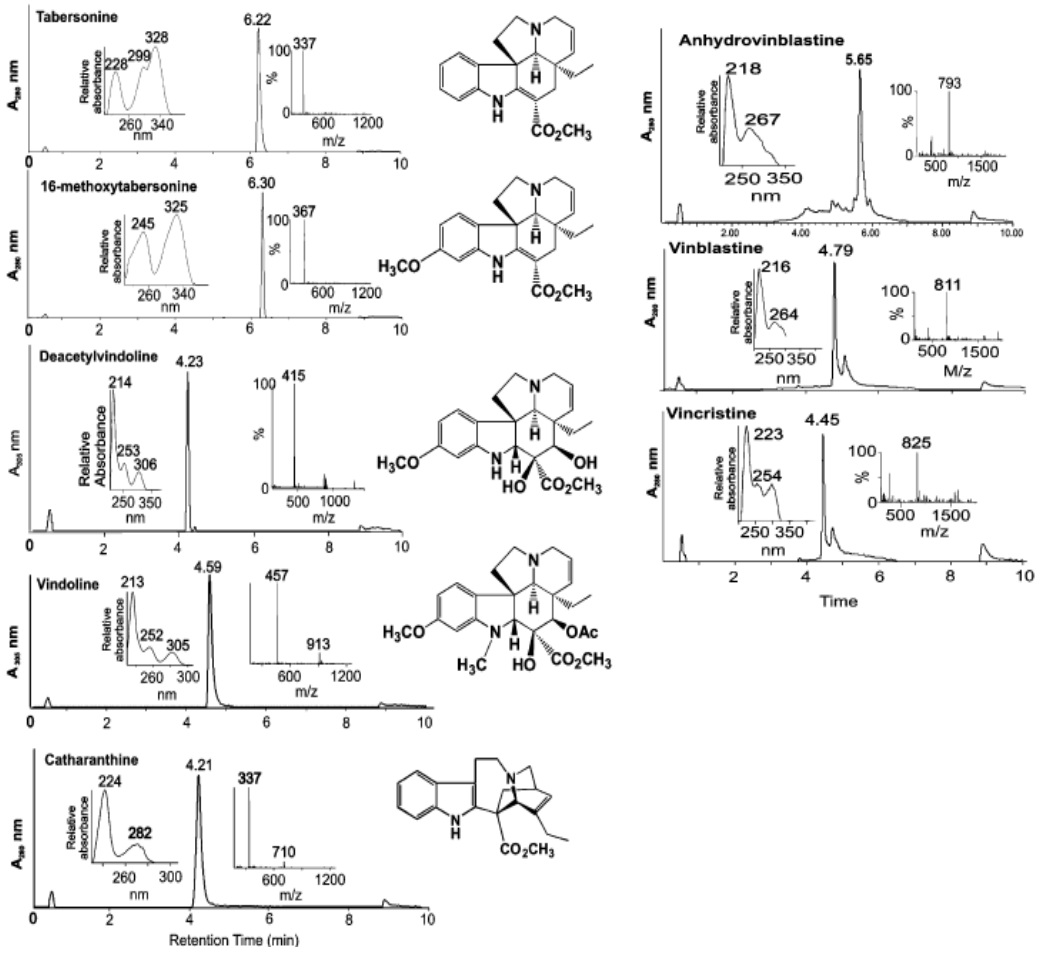


Figure 2.S1 UPLC-DAD-MS profiles of various MIAs from *C. roseus*: tabersonine, 16-methoxytabersonine, deacetylvindoline, vindoline, catharanthine, anhydrovinblastine, vinblastine, and vincristine. The retention time of each standard is indicated above each absorption peak.

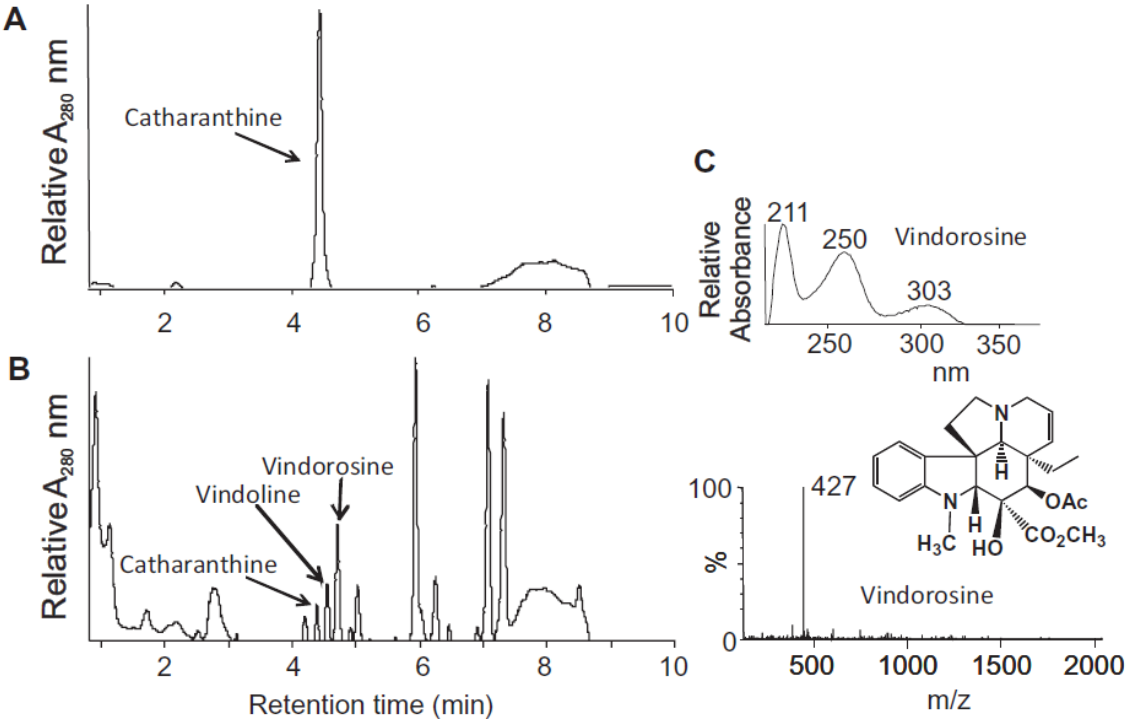


Figure 2.S2 Identification of catharanthine, vindoline, and vindorosine. MIAs extracted in chloroform after incubation of *C. roseus* (Mediterranean Deep orchid) leaves for 1 h were separated and identified and their concentrations were measured by UPLC-MS (*Upper*). After incubation in chloroform for 1 h, MIAs from leaves were extracted, separated, and identified, and their concentrations were measured by UPLC-MS (*Lower*). The peaks of catharanthine, vindoline, and vindorosine were identified by their retention times, their absorption spectra their respective masses. The absorption and mass spectra of vindorosine are displayed in Figure 2.S2C.

Appendix II

Chapter 3 Supplementary Tables:

Table 3.S1 PSPG-derived contigs found by searching EST database of *C. roseus*

PUT ID	Length	Top hit of Blast X	Top hit of Arabidopsis Blast X
Cr316	780	UDP-Glycosyltransferase (*Lycium barbatum*)	AtUGT85A2
Cr9886	940	UDP-Glycosyltransferase (*Vigna angularis*)	AtUGT85A1
Cr205	737	UDP-Glycosyltransferase (*Pisum sativum*)	AtUGT85A7
Cr31009	710	UDP-Glycosyltransferase family 1 protein (*Citrus sinensis*)	AtUGT85A5
Cr441	704	UDP-Glycosyltransferase family 1 protein (*Citrus sinensis*)	AtUGT85A7
Cr86	593	UDP-Glycosyltransferase,putative (*Ricinus communis*)	AtUGT85A7
Cr8440	596	UDP-Glycosyltransferase family 1 protein (*Citrus sinensis*)	AtUGT85A5
Cr9571	397	UDP-Glycosyltransferase,putative (*Ricinus communis*)	AtUGT85A1

Table 3.S2 Accession numbers of PSPGs used for constructing the phylogenetic tree shown in Figure 3.S1

Name	Species	Functions	Accession No.
AcGaT	*Aralia corbata*	Anthocyanin 3-*O*-galactosyltransferase	AB103471
B5GT	*Dorotheanthus*	Betanidin 5-*O*-glucosyltransferase	Y18871
Bx8	*Zea mays*	Benzoxazinoid 2-*O*-glucosyltransferase	AAL57037
BpUGAT	*Bellis perennis*	Anthocyanidin 3-*O*-glucoside 1,2-*O*-glucosyltransferase	BAD77944
CaUGT2	*Catharanthus roseus*	Curcumin glucosyltransferase	AB159213
CrUGT6（UGT85A23）	*Catharanthus roseus*	Iridoid 1-*O*-glucosyltransferase	AB591741
CrUGT7（UGT76A2）	*Catharanthus roseus*	Glucosyltransferase with broad substrate specificity	AB733666
CrUGT8（UGT709C2）	*Catharanthus roseus*	Iridoid 1-*O*-glucosyltransferase	AB733667
Cm1 2RhaT	*Citrus maxima*	Flavanone 7-*O*-glucoside 1,2-*O*-rhamnosyltransferase	AAL06646
GjUGT2（UGT85A24）	*Gardenia jasminoides*	Iridoid 1-*O*-glucosyltransferase	AB555732
F7GT	*Scutellaria baicalensis*	Flavonoid 7-*O*-glucosyltransferase	BAA83484
In3GGT	*Ipomoea nil*	Anthocyanidin 3-*O*-glucoside 2”-*O*-glucosyltransferase	AB192314
LjUGT12（UGT709C3）	*Lonicera japonica*	Iridoid 1-*O*-glucosyltransferase	
NtGT1a	*Nicotina tabacum*	Glucosyltransferase with broad substrate specificity	AB052558
NtGT2	*Nicotina tabacum*	Flavonol 7-*O*- coumarin 3-*O*-glucosyltransferase	AB072919
NtGT3	*Nicotina tabacum*	Glucosyltransferase with broad substarate specificity	AB072918
NtSAGT	*Nicotina tabacum*	Salicylic acid glucosyltransferase	AAF61647
6RhaT	*Petunia hybrida*	Anthocyanidin 3-*O*-glucoside 1,6-*O*-rhamnosyltransferase	CAA50376
SaGT4	*Solanum aculeatissimum*	Seroidal sapogenin 3-*O*-glucosyltransferase	BAD89042
UGT94D1	*Sesamum indicum*	Sesaminol 2-*O*-glucoside 1,6-*O*-glucosyltransferase	BAF99027
UGT85B1	*Sorghum bicolor*	*p*-Hydroxymandelonitrile *O*-glucosyltransferase	AAF17077
UGT76G1	*Stevia rebaudiana*	Stevioside glucosyltransferase	AAR06912
UGTCs2	*Crocus sativus*	Crocetin glucosyltransgerase	AAP94878
UGT71G1	*Medicago truncatula*	Flavonoid 3’-*O*-glucosyltransferase	AAW56092
UGT71C1	*Arabidopsis thaliana*	Hydroxycinamate, flavonoid 3-*O*-glucosyltransferase	AC005496
UGT72B1	*Arabidopsis thaliana*	Hydroxybenzoic acid glucosyltransferase	CAB80916
UGT72E2	*Arabidopsis thaliana*	Oniferylalcohol glucosyltransferase	AB018119
UGT73C5	*Arabidopsis thaliana*	Zeatin *O*-glucosyltransferase	AAD20156

Name	Species	Functions	Accession No.
UGT74B1	*Arabidopsis thaliana*	Thiohydroximate *S*-glucosyltransferase	AC002396
UGT75B1	*Arabidopsis thaliana*	Indole 3-acetic acid, sinapic acid glucosyltransferase	AAF79370
UGT76C1	*Arabidopsis thaliana*	*Trans*-zeatin *N*-glucosyltransferase	BAB10792
UGT78D1	*Arabidopsis thaliana*	Flavonol 3-*O*-glicoside *L*-rhamnosyltransferase	AC009917
UGT85A1	*Arabidopsis thaliana*	*Trans*-zeatin *O*-glucosyltransferase	AAF18537
VvGT1	*Vitis vinifera*	Flavonoid 3-*O*-glucosyltransferase	AAB81682
ZOG1	*Phaseolus lunatus*	Zeatin *O*-glucosyltransferase	AF101972
ZOX1	*Phaseolus lunatus*	Zeatin *O*-xylosyltransferase	AF116858
LjUGT8	*Lonicera japonica*	7-deoxyloganetic acid glucosyltransferase-like	KF415119
ClUGT8	*Cinchona ledgeriana*	7-deoxyloganetic acid glucosyltransferase-like	KF415120
AhUGT8	*Amsonia hubrichtii*	7-deoxyloganetic acid glucosyltransferase-like	KF415121
RsUGT8	*Rauwolfia serpentina*	7-deoxyloganetic acid glucosyltransferase-like	KF415122
TeUGT8	*Tabernaemontana elegans*	7-deoxyloganetic acid glucosyltransferase-like	KF415123
VmUGT8	*Vinca minor*	7-deoxyloganetic acid glucosyltransferase-like	KF415124
CloUGT8	*Catharanthus longifolius*	7-deoxyloganetic acid glucosyltransferase-like	KF415125
CovUGT8	*Catharanthus ovalis*	7-deoxyloganetic acid glucosyltransferase-like	KF415126

Table 3.S3 Primer sequences used for RACE-PCR cloning of *CrUGT6-8*

Primer name		Sequence（5'→3'）
CrUGT6	G7-Rv1 (5'-RACE first PCR)	CCACGAATTCTTCCCCAAGAATC
	G7-Rv2 (5'-RACE nested PCR)	CCAAGAATCGATGATTCGCAAGAG
CrUGT7	CrUGT7-5'first (5'-RACE first PCR)	CAAGGCCTACTTTCCATACATGAGT
	CrUGT7-5'nest (5'-RACE nested PCR)	GCTTCACAAATACTCTCAAGTGTTGAA
CrUGT8	Cr8440-Fw2 (3'-RACE PCR)	TGGTGACATTCCGATTAAAGGTAC
	Cr8440-Rv5'-5 (5'-RACE PCR)	AGCGACTGTGGCTGAACTCTG

Table 3.S4 Primer sequences used for full-length amplification of *CrUGT6-8*

Primer name		Sequence（5'→3'）
CrUGT6	G7-15	ATGGGTTCACTTTCTTCTTCCGATTACT
	G7	TCAATGACTTGGTTTTGATTTGAGCAC
CrUGT7	CrUGT7-Fw	ATGGGGTCCGTTTCTGCTAAA
	CrUGT7-Rv	TCTGCTATGGGAATGAGGCGAT
CrUGT8	Cr8440-head	ATGGGTTCTCAAGAAACAAATTTGGC
	Cr8440-Rv	GAAGTTTCTTTCTCTTCGTTGGTC

Table 3.S5 Primer sequences used for heterologous expression of *CrUGT6-8*

Primer name		Sequence (5'→3')
CrUGT6	G7-15-BamHI	GCGGATCCATGGGTTCACTTTCTTCTTCCGATTACT
	G7-SacI	GCGAGCTCTCAATGACTTGGTTTTGATTTGAGCAC
CrUGT7	CrUGT7-Fw-BamHI	GCGGATCCATGGGGTCCGTTTCTGCTAAA
	CrUGT7-Rv-SacI	GCGAGCTCCTATGGGAATGAGGCGAT
CrUGT8	CrUGT8-Fw-SphI	GCGCATGCATGGGTTCTCAAGAAACAAATTTGGC
	CrUGT8-Rv-SalI	GCGTCGACTCAAATAATCAGTGATTTTATGTAATCAACGAGG

Sequences shown in underline correspond to restriction enzyme sites that facilitated cloning.

Table 3.S6 Primer sequences used for real-time qRT-PCR

Primer name		Sequence (5'→3')
RPP0C (control) Reference gene	CrRPP0C-Fw-forRT	CAAGGTTGGAGCCCCTGCTCGTGTT
	CrRPP0C-RV-forRT	CAGGTGCTGAACATTCCAACTAAGA
CrUGT6	CrUGT6-for-RTFw2	CCATATCCAGCTCAAGGCCAC
	CrUGT6-for-RTRv2	GCCTTTGTAATGAAGGAGTTTTGC
CrUGT7	CrUGT7-for-RTFw1	CACATGACACCAATGCTTCAACTG
	CrUGT7-for-RTRv1	GATGAATTCAGGATGGTCTGAAGG
CrUGT8	CrUGT8-for-RTFw1	CCATGCTCAGACTAGCAGAAC
	CrUGT8-for-RTRv1	CAGAGTTCAGCCACAGTCGCTT
LAMT	LAMT-RT-Fw	GAAATGCCTGCTCTTCCAAC
	LAMT-RT-Rv	GTGGGAGTCATCACCACCTT
SLS	SLS-RT-Fw	CTTTGAGGGTGCAAAATGGT
	SLS-RT-Rv	TGGGATCCTTGTTTTTCAGC

Table 3.S7 Primer sequences used for VIGS in cloning partial sequences of CrUGT8, LAMT, SLS, and PDS, and in detection of TRV coat protein to confirm the success of *Agrobacterium* infiltration

Primer name		Sequence (5'→3')
CrUGT8	CrUGT8-VIGS-Fw	GACCAAATGATAAACAGCAG
	CrUGT8-VIGS-Rv	TTTTTGGTATCAACCCTAAT
LAMT	LAMT-VIGS-Fw	CTACACCAAAAATTTGCAAG
	LAMT-VIGS-Rv	TTGATTAAGGTCCTCTTGACTAG
SLS	SLS-VIGS-Fw	AGAGCATTTGGGGAGAAGAT
	SLS-VIGS-Rv	AGGATTCAGCTTGGAGTGCT
PDS	PDS-VIGS-Fw	AGGTTTGGGGGGTTTGTGT
	PDS-VIGS-Rv	TACGCCTTGCTTTCTCATCC
TRV coat protein	TRVCP-Fw	CGGGCTAACAGTGCTCTTG
	TRVCP-Rv	CTCCCTTGGTTCGTCGTAAC

Chapter 3 Supplementary Figures:

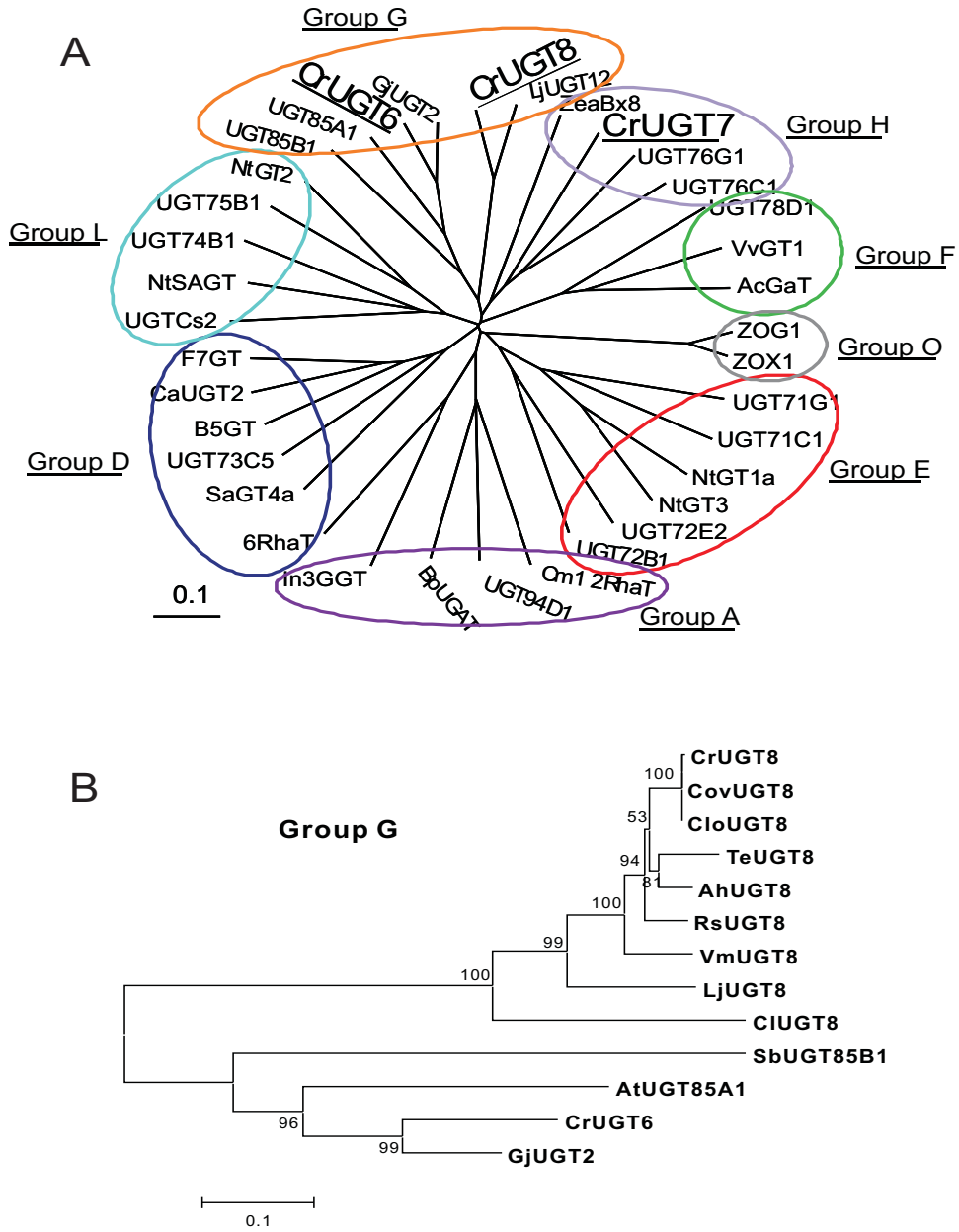


Figure 3.S1 Non-rooted molecular phylogenetic trees of plant UGTs. A) *Catharanthus roseus CrUGT6*, *CrUGT7* and *CrUGT6* described in the present investigation are located in group G of Family 1 GTs. The names and DDBJ/GenBank/EBI accession numbers of UGTs used for the alignment are shown in Supplementary Table 3.S1. **B)** Phylogenetic relationship among CrUGT8 genes in MIA producing members of the Apocynaceae family and in a secologanin producing member of the Caprifoliaceae. The numbers on each branch represent bootstrap values for 10000 replicates. Abbreviations: Cr-*Catharanthus roseus*, Vm-*Vinca minor*, Rs-*Rauwolfia serpentina*, Te-*Tabernaemontana elegans*, Ah-*Amsonia hubrichtii*, Cl-*Cinchona ledgeriana*, Lj-*Lonicera japonica*, Gj-*Gardenia jasminoides*, Sb-*Sorghum bicolor*, At-*Arabidopsis thaliana*, Clo-*Catharanthus longifolius*, Cov-*Catharanthus ovalis* using the neighbor-joining algorithm. The trees were constructed from a CLUSTALW program multiple alignment using the neighbor-joining method of MEGA 5. *Bar* = 0.1 amino acid substitutions/site.

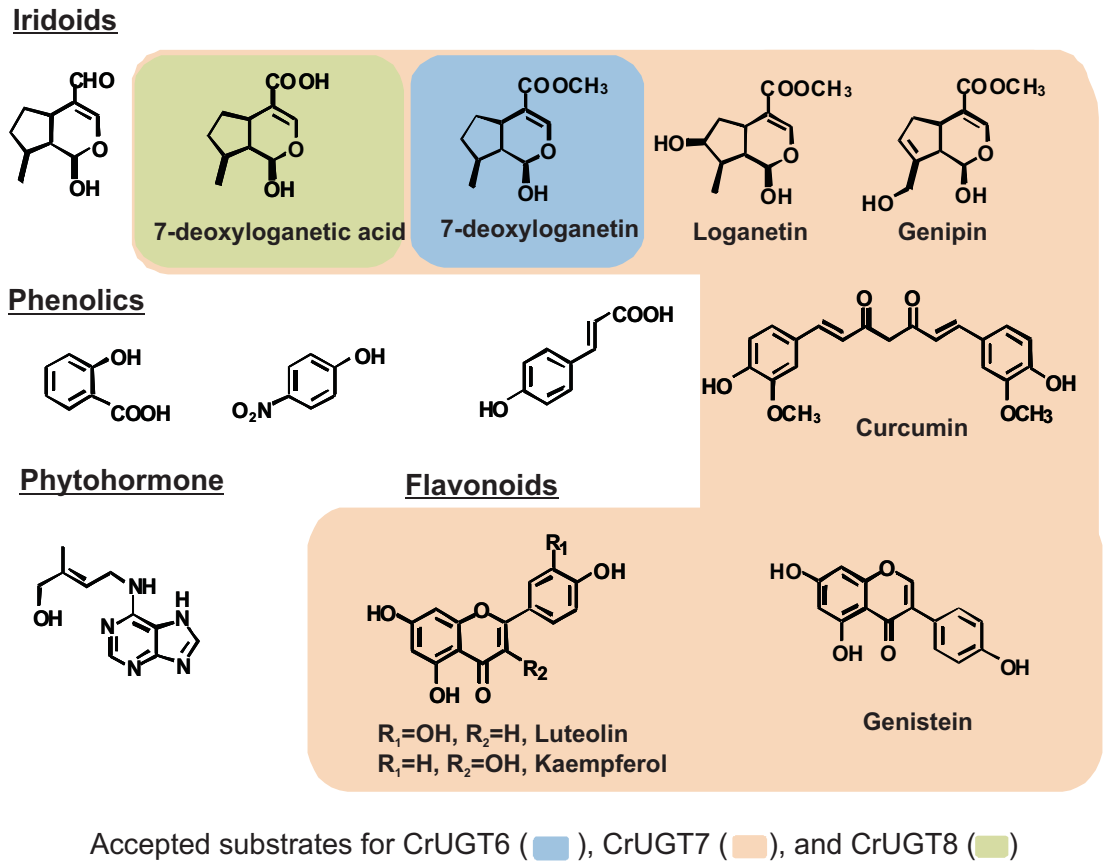


Figure 3.S2 Substrate specificity of recombinant CrUGT6, CrUGT7, and CrUGT8

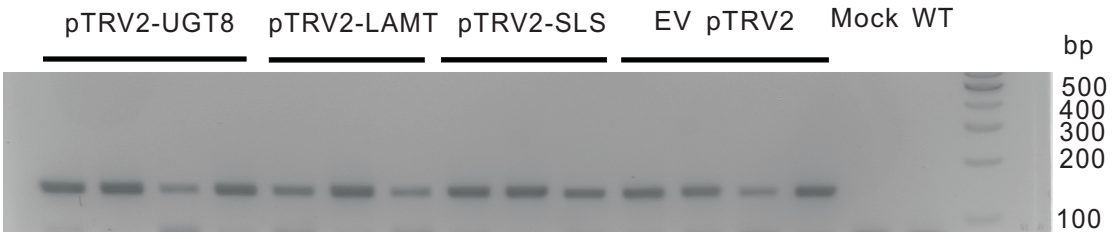


Figure 3.S3 Detection of pTRV2-derived TRV coat protein transcript (134 bp) in plants infiltrated with pTRVs vectors confirmed the success of *Agrobacterium* infiltration. Neither wild type plants that were not infiltrated nor mock-treated plants showed the TRV coat protein transcript.

Appendix III

Chapter 4 Supplementary Tables:

Table 4.S1 Cytochrome P450 genes in *C. roseus* retrieved from PhytoMetaSyn database generated based on 454 large-scale sequencing that contains 26,804 sequences in the order of their abundance

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
SLS	cytochrome P450 [*Catharanthus roseus*]	gi|404688|	0.0	1032	Iridoid	72	✓
170	cytochrome P450 [*Catharanthus roseus*]	gi|365927738|	0.0	417	Early MIA	71	✓
316	PREDICTED: cytochrome P450 76A2 [*Vitis vinifera*]	gi|225441018|	0.0	242	Iridoid	76	-
611	cytochrome P450 [*Catharanthus roseus*]	gi|365927740|	0.0	174	Late MIA	71	✓
479	cytochrome P450 [*Catharanthus roseus*]	gi|365927742|	0.0	166	Triterpene	716	✓
DL7H	putative secologanin synthase [*Camptotheca acuminata*]	gi|315439540|	0.0	108	Iridoid	72	-
HPL	hydroperoxide lyase [*Nicotiana attenuata*]	gi|15982240|	0.0	103	Fatty acid	74	✓
905	cytochrome P450, putative [*Ricinus communis*]	gi|255580537|	0.0	102	Late MIA	82	-
G10H	geraniol 10-hydroxylase [*Catharanthus roseus*]	gi|17065916|	0.0	91	Iridoid	76	-
885	tabersonine 16-hydroxylase CYP71D12 [*Catharanthus roseus*]	gi|223587559|	0.0	88	Late MIA	71	✓
1685	cytochrome P450 [*Catharanthus roseus*]	gi|325989355|	0.0	86	Late MIA	81	-
11807	Cytochrome P450 [*Medicago truncatula*]	gi|357494829|	1e-128	85	Late MIA	71	-
202	PREDICTED: cytochrome P450 77A2 [*Vitis vinifera*]	gi|225449736|	0.0	83	Non- MIA	77	✓

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
C4H	RecName: Full=Trans-cinnamate 4-monooxygenase	gi|1351206|	0.0	80	Phenyl-propanoid	73	✓
T16H	tabersonine 16-hydroxylase CYP71D12 [*Catharanthus roseus*]	gi|223587559|	0.0	79	Vindoline	71	✓
1881	obtusifoliol-14-demethylase [*Petunia x hybrida*]	gi|118200120|	0.0	76	Sterol	51	-
1663	cytochrome P450 [*Catharanthus roseus*]	gi|325989355|	0.0	75	Late MIA	81	-
13535	4-coumaric acid 3`-hydroxylase 34 [*Coffea arabica*]	gi|399630574|	0.0	73	Phenyl-propanoid	98	-
3969	cytochrome P450 [*Centaurium erythraea*]	gi|46403213|	2e-107	70	Late MIA	76	-
1700	cytochrome P450 [*Catharanthus roseus*]	gi|12657333|	0.0	62	Non- MIA	71	✓
1935	PREDICTED: cytochrome P450 71D10-like [*Vitis vinifera*]	gi|359494295|	0.0	59	Late MIA	71	-
2463	cytochrome P450-type monooxygenase 97A29 [*Solanum lycopersicum*]	gi|350537695|	0.0	56	Non- MIA	97	-
23379	PREDICTED: protein [*Populus trichocarpa*]	gi|224125112|	0.0	52	Non- MIA	93	✓
1834	CYP72A54 [*Nicotiana tabacum*]	gi|85068674|	0.0	51	Iridoid	72	-
2039	PREDICTED: cytochrome P450 82A3-like [*Vitis vinifera*]	gi|359492643|	0.0	50	Late MIA	82	-
1933	hypothetical protein VITISV_037194 [*Vitis vinifera*]	gi|147844260|	0.0	48	Non- MIA	86	-
4892	PREDICTED: cytochrome P450 97B2, chloroplastic [*Vitis vinifera*]	gi|225464788|	0.0	35	Non- MIA	97	-
4334	cytochrome P450 90A2 [*Camellia japonica*]	gi|71725821|	0.0	35	Non- MIA	90	-
3732	PREDICTED: flavonoid 3'-monooxygenase-like [*Vitis vinifera*]	gi|225442104|	0.0	32	Flavonoid	92	-

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
4675	cytochrome P450 [*Ipomoea nil*]	gi|124484371|	0.0	32	Late MIA	86	-
2515	cytochrome P450 [*Nicotiana tabacum*]	gi|252972629|	0.0	32	Late MIA	81	-
4494	cytochrome P450, putative [*Ricinus communis*]	gi|255566096|	0.0	29	Non- MIA	89	-
3342	PREDICTED: cytochrome P450 94A1 [*Vitis vinifera*]	gi|225440161|	0.0	27	Non- MIA	94	-
9374	CYP81B2v2 [*Nicotiana tabacum*]	gi|85068594|	0.0	25	Early MIA	81	✓
5085	PREDICTED: allene oxide synthase, chloroplastic [*Vitis vinifera*]	gi|225428606|	0.0	25	Fatty acid	74	-
4610	PREDICTED: cytochrome P450 734A1 [*Vitis vinifera*]	gi|225448341|	0.0	24	Non- MIA	714	-
7400	4-coumaric acid 3`-hydroxylase 25 [*Coffea arabica*]	gi|399630550|	0.0	22	Phenyl-propanoid	98	✓
5782	cytochrome P450 [*Catharanthus roseus*]	gi|325989355|	4e-127	22	Non- MIA	81	-
6791	CYP72A57 [*Nicotiana tabacum*]	gi|85068668|	0.0	19	Non- MIA	72	-
6776	CYP72A57 [*Nicotiana tabacum*]	gi|85068668|	0.0	19	Early MIA	72	-
5944	cytochrome P450, putative [*Ricinus communis*]	gi|255563438|	0.0	19	Non- MIA	71	-
10559	cytochrome P450 CYP85A1 [*Nicotiana tabacum*]	gi|109649528|	0.0	19	Non- MIA	85	-
8899	cytochrome P450 [*Catharanthus roseus*]	gi|12657333|	0.0	18	Late MIA	71	-
5019	PREDICTED: cytochrome P450 81F1-like [*Vitis vinifera*]	gi|225458753|	1e-174	17	Late MIA	81	-
7301	PREDICTED: cytochrome P450 716B2-like [*Vitis vinifera*]	gi|359478809|	2e-162	17	Late MIA	716	-

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
10317	cytochrome P450 CYP736A54 [*Bupleurum chinense*]	gi|388827897|	5e-173	16	Non- MIA	71	-
25640	Ent-kaurenoic acid oxidase putative [*Ricinus communis*]	gi|255571147|	0.0	15	Gibberellin	88	-
6110	cytochrome P450 [*Populus trichocarpa*]	gi|224133724|	1e-175	14	Non- MIA	71	-
8832	unnamed protein product [*Vitis vinifera*]	gi|297740004|	0.0	14	Non- MIA	97	-
5869	allene oxide synthase [*Solanum tuberosum*]	gi|20160362|	0.0	13	Fatty acid	74	-
10679	coniferylalcohol 5-hydroxylase [*Centaurium erythraea*]	gi|46403211|	0.0	12	Non- MIA	84	✓
3946	predicted protein [*Populus trichocarpa*]	gi|224145360|	3e-53	10	Non- MIA	81	-
8109	cytochrome P450 [*Populus trichocarpa*]	gi|224131024|	0.0	10	Non- MIA	734	✓
10250	cytochrome P450-dependent fatty acid hydroxylase [*Nicotiana tabacum*]	gi|18000072|	0.0	10	Fatty acid	94	-
13060	cytochrome P450 monooxygenase isoform I [*Sesamum indicum*]	gi|70724310|	1e-107	9	Late MIA	76	-
14926	PREDICTED: cytochrome P450 82C4-like [*Vitis vinifera*]	gi|359497458|	0.0	9	Non- MIA	82	-
10517	ABA 8'-hydroxylase [*Citrus sinensis*]	gi|367465456|	0.0	9	Abscisic acid	707	-
16667	hypothetical protein VITISV_035274 [*Vitis vinifera*]	gi|147767047|	0.0	8	Non- MIA	71	✓
9207	cytochrome P450 monooxygenase isoform I [*Sesamum indicum*]	gi|70724310|	0.0	8	Late MIA	76	-
5607	PREDICTED: isoflavone 2'-hydroxylase-like [*Vitis vinifera*]	gi|225458768|	4e-41	7	Non- MIA	81	-

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
8541	PREDICTED: isoflavone 2'-hydroxylase-like [*Vitis vinifera*]	gi|359491991|	4e-53	7	Non- MIA	81	-
12089	cytochrome P450 monooxygenase [*Petunia x hybrida*]	gi|71726950|	0.0	7	Non- MIA	704	-
24314	hypothetical protein 111O18.23 [*Coffea canephora*]	gi|338762849|	0.0	6	Non- MIA	90	-
16781	cytochrome P450 [*Populus trichocarpa*]	gi|224122436|	4e-133	5	Late MIA	76	-
16109	cytochrome P450 [*Ipomoea nil*]	gi|124484371|	0.0	5	Non- MIA	86	-
12228	PREDICTED: isoflavone 2'-hydroxylase-like [*Vitis vinifera*]	gi|359480637|	3e-56	5	Non- MIA	81	-
18596	cytochrome P450, putative [*Ricinus communis*]	gi|255585725|	0.0	5	Non- MIA	94	-
17333	PREDICTED: cytochrome P450 81D1-like [*Glycine max*]	gi|356540926|	1e-28	4	Non- MIA	81	-
17434	PREDICTED: flavonoid 3'-monooxygenase [*Vitis vinifera*]	gi|225457235|	0.0	4	Flavonoid	92	-
16069	Cytochrome P450 [*Medicago truncatula*]	gi|357494853|	0.0	3	Late MIA	71	-
18386	cytochrome P450 [*Catharanthus roseus*]	gi|365927740|	7e-31	3	Non- MIA	71	✓
15759	PREDICTED: isoflavone 2'-hydroxylase [*Vitis vinifera*]	gi|359480598|	1e-45	3	Non- MIA	81	-
18493	cytochrome P450 71 family protein [*Solanum lycopersicum*]	gi|350539956|	1e-77	3	Non- MIA	71	-
19339	Premnaspirodiene oxygenase	gi|334305730|	9e-111	3	Non- MIA	71	-
15980	PREDICTED: cytochrome P450 85A [*Vitis vinifera*]	gi|225423426|	0.0	3	Non- MIA	85	-

Gene ID	Description	Hit ID	E value	# EST	Predicted pathway	Family	Enriched in epidermis
19795	cytochrome P450 [*Populus trichocarpa*]	gi|224136932|	2e-63	2	Non- MIA	76	-
20991	cytochrome P450, putative [*Ricinus communis*]	gi|255547149|	0.0	2	Non- MIA	78	-
19333	cytochrome P450 monooxygenase CYP72A59 [*Medicago truncatula*]	gi|84514139|	8e-33	2	Non- MIA	72	-
19375	sterol 22-desaturase [*Solanum lycopersicum*]	gi|350536665|	0.0	2	Fatty acid	710	-
20988	cytochrome P450 [*Catharanthus roseus*]	gi|325989355|	7e-51	2	Non- MIA	81	-
19096	tabersonine 16-hydroxylase CYP71D12 [*Catharanthus roseus*]	gi|223587559|	4e-36	2	Late MIA	71	-
26447	RecName: Full=Cytochrome P450 71A8	gi|5915815|	3e-30	2	Non- MIA	71	-
20699	cytochrome P450 [*Populus trichocarpa*]	gi|224130624|	8e-54	2	Non- MIA	78	-
23212	PREDICTED: cytochrome P450 81D1-like [*Glycine max*]	gi|356540926|	0.0	2	Non- MIA	81	-
19753	tabersonine/lochnericine 19-hydroxylase [Catharanthus roseus]	gi|325989353|	9e-91	2	Late MIA	71	-
22048	PREDICTED: 3-epi-6-deoxocathasterone 23-monooxygenase [*Vitis vinifera*]	gi|225431255|	0.0	2	Non- MIA	90	-
1002	cytochrome P450 [*Populus trichocarpa*]	gi|224111880|	0.0	1	Non- MIA	734	-

Genes that are shown in bold characters are the clones that have been functionally characterized in *C. roseus*. "✓" indicates the enrichment of the genes in the epidermis database (http:// NapGen (Natural Products Genomic Research) (Murata *et al*., 2006). Light grey shading indicates the gene being characterized in the present study.

Table 4.S2 Genes in Apocynaceae family, *Cinchona ledgeriana*, and *Lonicera japonica* with their percent similarity to known MIA biosynthetic genes in *C. roseus*

Catharanthus roseus	**Enriched in epidermis**	***Lonicera japonica***	***Cinchona ledgeriana***	***Amsonia hubrichtii***	***Rauwolfia serpentina***	***Tabernae-montana elegans***	***Vinca minor***
GES	-	82	74	83	79	84	81
G10H	-	85	84	86	88	87	86
10HGO	-	78	78	89	90	89	90
IRS	-	61	83	88	89	88	90
CrUGT8	-	76	83	91	92	90	87
DL7H	-	74	81	90	92	90	89
LAMT	✓	70	81	85	85	85	85
SLS	✓	76	77	92	93	91	91
TDC	✓	67	71	84	86	84	88
STR	✓	-	57	78	82	71	79
SGD	✓	-	52	52	70	70	72
T16H	✓	-	-	-	-	-	-
16OMT	✓	-	-	-	-	-	-
NMT	-	-	-	-	-	-	-
D4H	-	-	-	-	-	-	-
DAT	-	-	-	-	-	-	-

Grey shading indicates iridoid biosynthetic genes and highlighted row is the gene being investigated in the present study. "-" shows that the ratio of enrichment epidermis/whole leaf calculated as previously described in De Luca *et al*., 2012b is lower than 1.5 as cut-off value to suggest no enrichment in epidermis. "✓" indicates the enrichment in the epidermis (Ratio of enrichment epidermis over whole leaf > 1.5). The numbers across this table show the percent similarity of genes found in each species compared to MIA biosynthetic genes in *Catharanthus roseus*. The cut off values for the post-strictosidine biosynthetic genes were 50%, and for the vindoline biosynthetic genes were 72%. Abbreviations: GES-geraniol synthase; G10H-geraniol 10-hydroxylase; 10HGO-10-hydroxygeraniol oxidoreductase; IRS-iridoid synthase; CrUGT8-7-deoxyloganetic acid glucosyltransferase (Chapter 3); DL7H-7-deoxyloganic acid 7-hydroxylase; LAMT-loganic acid *O*-methyltransferase; SLS-secologanin synthase; TDC-tryptophan decarboxylase; STR-strictosidine synthase; SGD-strictosidine β-D-glucosidase; T16H-tabersonine 16-hydroxylase; 16OMT-16-hydroxytabersonine *O*-methyltransferase; NMT-16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase; D4H-deacetoxyvindoline 4-hydroxylase; DAT-deacetylvindoline 4-acetyltransferase.

Table 4.S3 Accession numbers of genes used for constructing the phylogenetic tree shown in Figure 4.7

Name	Species	Functions	Accession No.
CrDL7H	*Catharanthus roseus*	7-deoxyloganic acid 7-hydroxylase	KF415115
CrSLS	*Catharanthus roseus*	Secologanin synthase CYP72A1	L10081
LjSLS	*Lonicera japonica*	Secologanin synthase-like	KF415097
ClSLS	*Cinchona ledgeriana*	Secologanin synthase-like	KF415098
AhSLS	*Amsonia hubrichtii*	Secologanin synthase-like	KF415099
RsSLS	*Rauwolfia serpentina*	Secologanin synthase-like	KF415100
TeSLS	*Tabernaemontana elegans*	Secologanin synthase-like	KF415101
VmSLS	*Vinca minor*	Secologanin synthase-like	KF415102
CaSLS	*Camptotheca acuminata*	Secologanin synthase-like	HQ605982
CrG10H	*Catharanthus roseus*	Geraniol 10-hydroxylase CYP76B6	AJ251269
LjG10H	*Lonicera japonica*	Geraniol 10-hydroxylase-like	KF415103
ClG10H	*Cinchona ledgeriana*	Geraniol 10-hydroxylase-like	KF415104
AhG10H	*Amsonia hubrichtii*	Geraniol 10-hydroxylase-like	KF415105
RsG10H	*Rauwolfia serpentina*	Geraniol 10-hydroxylase-like	KF415106
TeG10H	*Tabernaemontana elegans*	Geraniol 10-hydroxylase-like	KF415107
VmG10H	*Vinca minor*	Geraniol 10-hydroxylase-like	KF415108
LjDL7H	*Lonicera japonica*	7-deoxyloganic acid 7-hydroxylase-like	KF415109
ClDL7H	*Cinchona ledgeriana*	7-deoxyloganic acid 7-hydroxylase-like	KF415110
AhDL7H	*Amsonia hubrichtii*	7-deoxyloganic acid 7-hydroxylase-like	KF415111
RsDL7H	*Rauwolfia serpentina*	7-deoxyloganic acid 7-hydroxylase-like	KF415112
TeDL7H	*Tabernaemontana elegans*	7-deoxyloganic acid 7-hydroxylase-like	KF415113
VmDL7H	*Vinca minor*	7-deoxyloganic acid 7-hydroxylase-like	KF415114
GuCYP72A154	*Glycyrrhiza uralensis*	C-30 oxidation of 11-oxo-β-amyrin	AB558153
MtCYP72A63	*Medicago truncatula*	C-30 oxidation of β-amyrin	AB558146
MtCYP72A61v2	*Medicago truncatula*	24-OH-β-amyrin to soyasapogenol	AB558145
MtCYP72A68v2	*Medicago truncatula*	oleanolic acid to gypsogenic acid	AB558150
CrC4H	*Catharanthus roseus*	Cinnamate 4-hydroxylase CYP73A4	Z32563
CrT19H	*Catharanthus roseus*	Tabersonine 19-hydroxylase CYP71BJ1	HQ901597
CrT16H	*Catharanthus roseus*	Tabersonine 16-hydroxylase CYP71D12	FJ647194
CrF3,5H	*Catharanthus roseus*	Flavonoid 3',5'-hydroxylase CYP75A8	AJ011862

Chapter 4 Supplementary Figures:

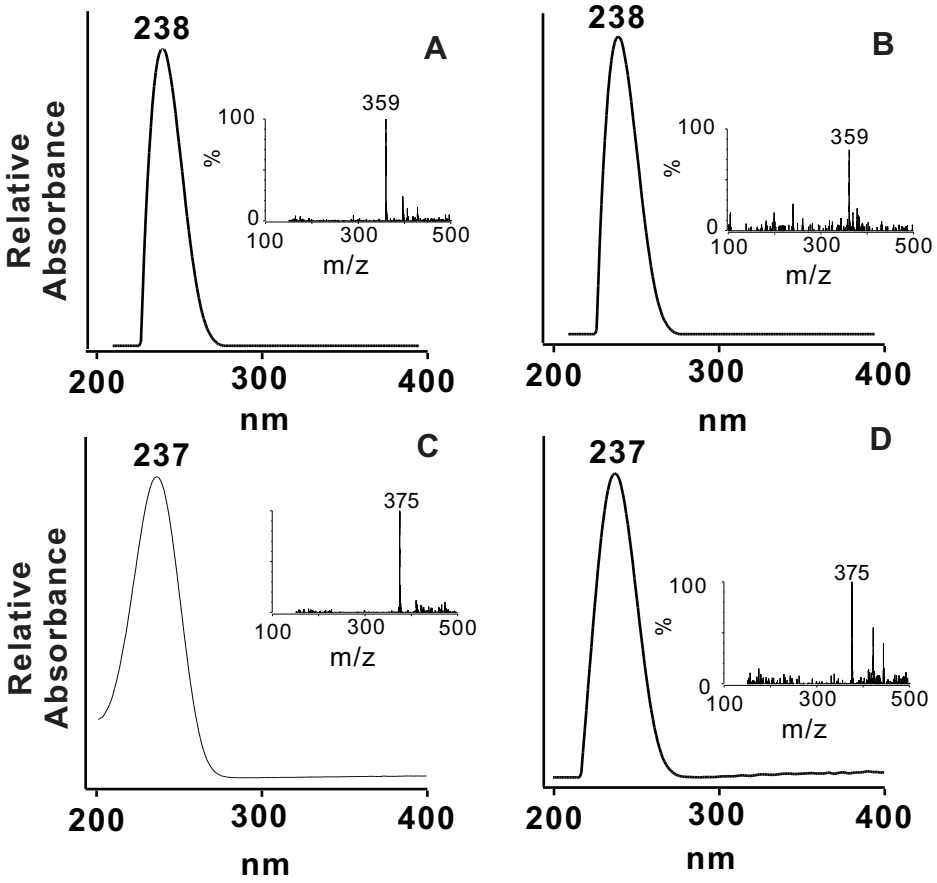


Figure 4.S1 Identification of 7-deoxyloganic acid in VIGS and loganic acid in *CrDL7H* enzyme assays. Representation of UV and mass spectrum of iridoid extracted from VIGS-treated plants at RT=4.06 min (B) was compared with authentic 7-deoxyloganic acid (A), m/z=359 and major absorbance at 238 nm. The UV and mass spectrum of products in enzyme assay (D) at RT=1.9 min was compared with authentic loganic acid (C), m/z=375 and major absorbance at 237 nm.

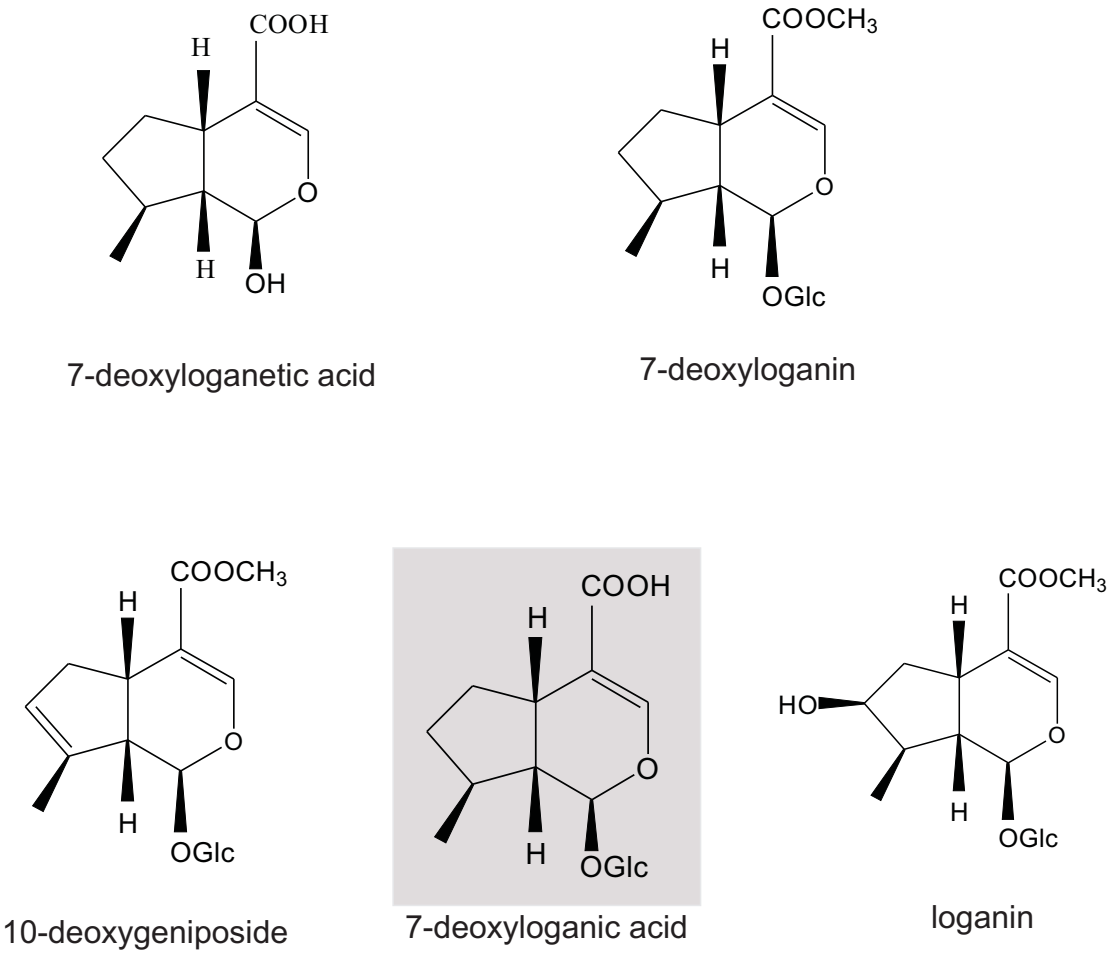


Figure 4.S2 Substrate specificity of recombinant *CrDL7H*. Structures of iridoids tested as substrates with *CrDL7H*. Only 7-deoxyloganic acid (highlighted in gray) was accepted as a substrate.

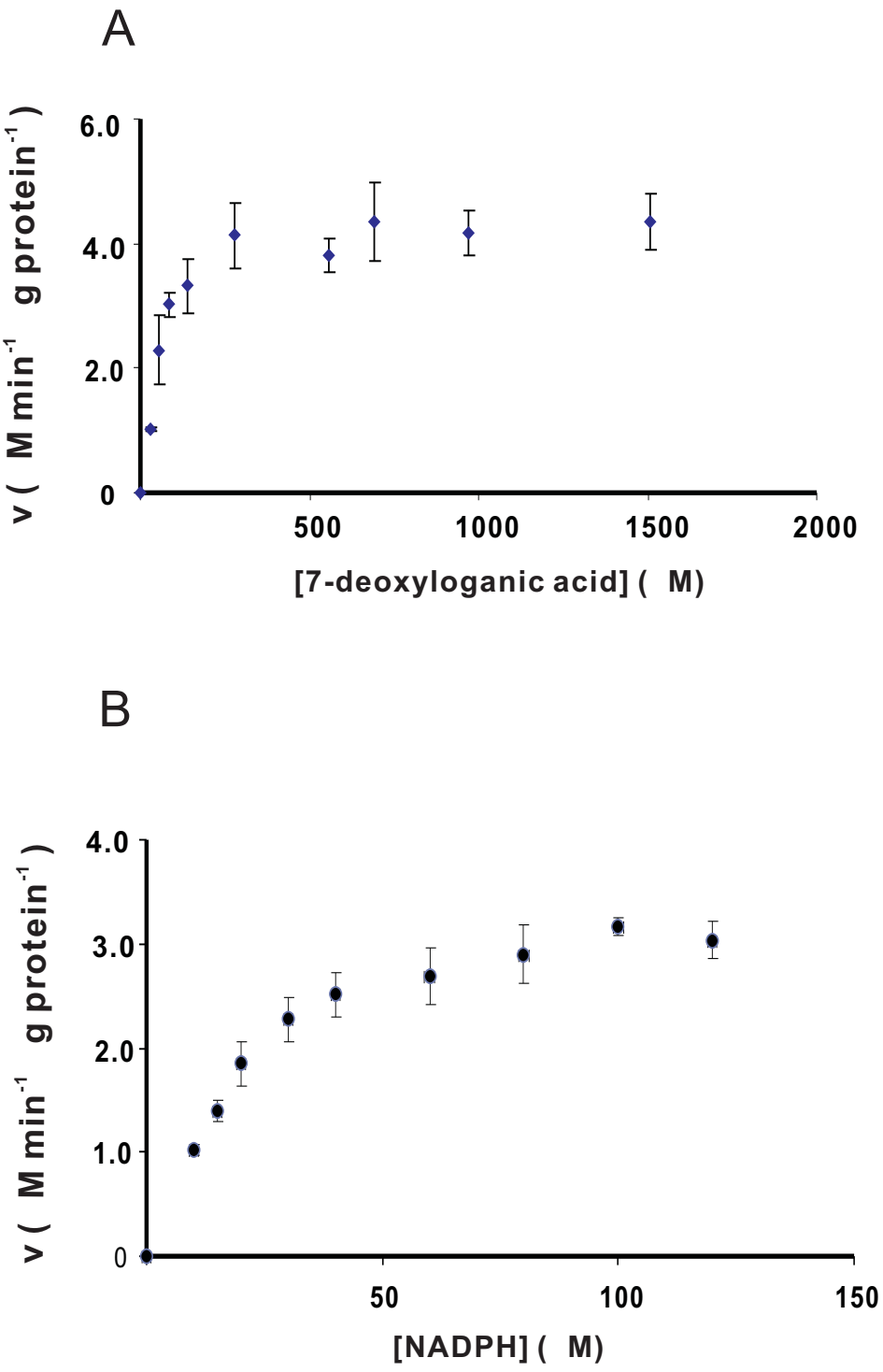


Figure 4.S3 Steady-state Michaelis-Menten kinetics derived from initial rates of *CrDL7H*-enriched microsomes with (**A**) 7-deoxyloganic acid and (**B**) NADPH. Error bars represent the 95% confidence of the standard deviations from three independent experiments.